Human-Centered and Computational Understanding for the Design and Adaptation of Mental Health and Well-being Interventions

Jina Suh

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

James Fogarty, Chair

Tim Althoff, Chair

Mary Czerwinski

Program Authorized to Offer Degree:
Computer Science & Engineering

University of Washington

Abstract

Human-Centered and Computational Understanding for the Design and Adaptation of Mental Health and Well-being Interventions

Jina Suh

Co-Chairs of the Supervisory Committee:
Professor James Fogarty
UW Computer Science & Engineering

Assistant Professor Tim Althoff
UW Computer Science & Engineering

As many as 20% of Americans suffer from diagnosable mental health disorders, but those overwhelmed with physiological and economic burdens cannot prioritize seeking support for their mental health and well-being. There are many evidence-based psychosocial interventions (EBPIs) that have been proven to be effective in treating mental health conditions. Recent initiatives to improve engagement in mental health care through technology have generated an abundance of promising digital mental health solutions. However, symptoms of stress, anxiety, and depression remain overlooked and in constant tension with life demands and disruptions, making it challenging to integrate such solutions into everyday life. My dissertation research examines the tensions between everyday life demands and mental health and well-being, where I design systems that integrate adaptations of EBPIs into everyday contexts to promote engagement. My work intersects three well-being contexts: (1) the COVID-19 pandemic, (2) co-morbid cancer and depression, and (3) workplace stress.

First, I examine the situated contexts using human-centered and computational methods grounded on holistic frameworks to reveal challenges rooted in tensions among multiple needs that get in the way of engaging in mental health and well-being activities. I conduct this research in the COVID-19 pandemic and co-morbid cancer and depression contexts to demonstrate that these challenges are present at the individual, organizational, and population scales. Second, I identify modification targets to existing evidence-based psychosocial interventions that can be enhanced through the use of technology to ease the tensions among needs and to directly integrate adapted interventions into the relevant contexts. I describe the development of the collaborative behavioral activation system aimed at improving the collaboration and engagement of patients and providers in depression care. I also describe the development of a just-in-time micro-intervention system aimed at reducing stress in the workplace. Lastly, I deploy these technology-enhanced mental health and well-being systems in real-world contexts to evaluate their effectiveness in improving engagement. Through such deployment, I highlight implementation challenges to integrating patient-provider collaborative technology into a clinical care practice as well as individual, contextual, and intervention-related factors that may influence real-time engagement in digitized interventions. Across three well-being contexts, my dissertation demonstrates that contextual and continuous adaptations of EBPIs can improve engagement in mental health and well-being care. My dissertation makes theoretical contributions through the development of holistic frameworks, methodological contributions through the development of computational frameworks, and artifact contributions through the development of technology-enhanced mental health and well-being intervention systems and through the design recommendations that arise from real-world deployments.

TABLE OF CONTENTS

Page

List of Figures . . . iv

List of Tables . . . vii

Chapter 1: Introduction . . . 1
  1.1 Thesis Statement . . . 3
  1.2 Dissertation Overview . . . 3
  1.3 Organization of the Dissertation . . . 7

Chapter 2: Related Work and Background . . . 9
  2.1 Mental Health and Well-being Interventions . . . 9
  2.2 Technology Support for Mental Health and Well-being . . . 11
  2.3 Mental Health and Well-being Contexts . . . 14
    2.3.1 The COVID-19 Pandemic . . . 14
    2.3.2 Cancer and Depression . . . 16
    2.3.3 Workplace Stress . . . 20

Chapter 3: Understanding the Impact of the Pandemic on Human Needs . . . 23
  3.1 Human Needs Framework and Shifting Priorities . . . 25
    3.1.1 Dataset and Validity . . . 28
    3.1.2 Human Needs Framework . . . 30
    3.1.3 Temporal Changes in Human Needs . . . 34
    3.1.4 Significant Changes in Human Needs . . . 36
    3.1.5 Geographical Differences in Needs . . . 40
    3.1.6 Gaps between Expressed and Reported Needs . . . 42
    3.1.7 Discussion . . . 45
  3.2 Disparate Impact and Burdens of the Pandemic . . . 47

3.2.1 Methods . . . 50
3.2.2 Results . . . 64
3.2.3 Discussion . . . 72
3.3 Summary of Contributions to Thesis . . . 76

Chapter 4: Supporting Collaborative Care for Patients with Cancer and Depression . . 79
4.1 Parallel Journeys Framework and Competing Demands . . . 81
4.1.1 Method . . . 83
4.1.2 Parallel Journeys Framework . . . 91
4.1.3 Two Journeys, Additional Challenges . . . 99
4.1.4 Technology Design Opportunities for Collaborative Care for Depression and Cancer . . . 110
4.1.5 Discussion and Future Work . . . 118
4.2 Design of Technology-Enhanced Collaborative Care . . . 121
4.2.1 Digitization of Collaborative Care and Behavioral Activation . . . 123
4.2.2 SCOPE System Design . . . 129
4.2.3 Study Design and Method . . . 136
4.2.4 Findings . . . 142
4.2.5 Discussion . . . 159
4.3 Summary of Contributions to Thesis . . . 168

Chapter 5: Integrating Stress-Reduction Interventions in the Workplace . . . 170
5.1 Providing Integrated Support for Stress Reduction at Work . . . 173
5.1.1 Method . . . 176
5.1.2 Findings . . . 192
5.1.3 Discussion . . . 201
5.2 Evaluation of Intervention Engagement at Work . . . 208
5.2.1 Motivating the Research Questions . . . 211
5.2.2 Data, Processing, and Analysis . . . 212
5.2.3 Results . . . 223
5.2.4 Discussion . . . 232
5.3 Summary of Contributions to Thesis . . . 240

Chapter 6: Discussion and Conclusion 243
6.1 Summary of Thesis Contributions 243
6.2 Design Recommendations and Future Directions 246
6.2.1 Consider the holistic context 246
6.2.2 Design adaptations for multiple temporal granularities 247
6.2.3 Navigate between personalization and standardization 249
6.2.4 Understand the role of technology through deployment in real-world contexts 250

Appendix A: Measuring Human Needs through Search Data 319
A.1 Analyzing National Representation 319
A.2 Analyzing Selection Bias 319
A.3 Reliability of Search Interaction Trends 320
A.4 Human Needs Categories and Detection 322
A.5 Human Need Detection Instructions for MTurk Workers 329
A.6 Significant Changes in Human Needs 333

Appendix B: Quantifying Digital Disparities through Search Data 334
B.1 SDoH factors used in analysis 334
B.2 Search categories 336
B.3 Unemployment claims 336
B.4 Matching Quality 337
B.5 Search trends 344
B.6 Relative differences across groups 355

Appendix C: Questionnaires for Stress Reduction Intervention Study 369
C.1 Ecological Momentary Assessments (EMAs) 369
C.2 Post-Study Questionnaire 370
C.2.1 Pre-scheduled condition 370
C.2.2 Just-in-time condition 372
C.2.3 Common 373

LIST OF FIGURES

Figure Number Page

1.1 Scope of dissertation research work and methods 6

2.1 Illustration of Collaborative Care Model 18

3.1 Illustration of human needs detection framework 28
3.2 Daily relative changes in needs during the pandemic 35
3.3 Daily relative changes in select needs during the pandemic 39
3.4 Relationships between shelter-in-place policies and needs during the pandemic . 42
3.5 Unemployment and domestic violence search behaviors during the pandemic . . . 43
3.6 Quantifying disparities in online health information access 52
3.7 Disparities in online economic assistance access 67
3.8 Disparities in shifting to digital resources 70

4.1 A snapshot of the psychosocial care journey mapping exercise 90
4.2 Four phases of a psychosocial care journey 93
4.3 Illustration of parallel journeys for three patients with cancer and depression . . . 96
4.4 Illustration of a modified Collaborative Care Model with technology-enhanced registry 125
4.5 Collaborative behavioral activation loop model 128
4.6 SCOPE system capabilities 130
4.7 Screenshots from the SCOPE registry 132
4.8 Screenshots from the SCOPE behavioral activation app 134

5.1 Browsing the intervention catalog 179
5.2 Huey’s intervention nudge 182
5.3 Intervention consumption dialog flow 183
5.4 Three intervention modalities supported by Huey 187
5.5 Study procedure across conditions 190
5.6 Self-reported stress levels throughout the study 193

5.7 Average momentary stress reduction and user rating 198
5.8 Average agreement on system usability questions 200
5.9 System-initiated intervention flow 210
5.10 Weekly trend of nudges, engagements, stress, and workload 218
5.11 Intervention nudge and engagement probabilities 220

A.1 Comparison of Bing and Google trends data on select keywords 322
A.2 Example labeling task for Amazon’s Mechanical Turk crowd worker 325
A.3 Daily relative changes in select needs during the pandemic 333

B.1 Our choice of census variables (red) and search categories (yellow) are inspired by the Social Determinants of Health framework (green). 334
B.2 Percent change in the unemployment related queries and clicks to unemployment sites 336
B.3 Normalized query volumes across five categories in years 2019 and 2020 across Bing and Google. 344
B.4 Frequency of search keywords “covid”, “coronavirus”, “mers”, “gonorrhea”, and “zika” in 2019 345
B.5 Query frequencies and proportions across groups with high or low percentages of Black or African American residents 346
B.6 Query frequencies and proportions across groups with high or low percentages of Hispanic residents 347
B.7 Query frequencies and proportions across groups with high or low median household income 348
B.8 Query frequencies and proportions across groups with high or low educational attainment 349
B.9 Query frequencies and proportions across groups with high or low population density 350
B.10 Query frequencies and proportions across groups with high or low unemployment rate 351
B.11 Query frequencies and proportions across groups with high or low health insurance coverage 352
B.12 Query frequencies and proportions across groups with high or low internet access 353
B.13 Percent change in ‘Health condition related queries’ between two matched groups across eight SDoH factors. 355

B.14 Differences in percentage points for changes in ‘Health condition related queries’ between two matched groups across eight SDoH factors. 356
B.15 Percent change in ‘Unemployment related queries’ between two matched groups across eight SDoH factors. 357
B.16 Differences in percentage points for changes in ‘Unemployment related queries’ between two matched groups across eight SDoH factors. 358
B.17 Percent change in ‘Clicks to state unemployment sites’ between two matched groups across eight SDoH factors. 359
B.18 Differences in percentage points for changes in ‘Clicks to state unemployment sites’ between two matched groups across eight SDoH factors. 360
B.19 Percent change in ‘Financial assistance related queries’ between two matched groups across eight SDoH factors. 361
B.20 Differences in percentage points for changes in ‘Financial assistance related queries’ between two matched groups across eight SDoH factors. 362
B.21 Percent change in ‘Click to online learning sites’ between two matched groups across eight SDoH factors. 363
B.22 Differences in percentage points for changes in ‘Click to online learning sites’ between two matched groups across eight SDoH factors. 364
B.23 Percent change in ‘Food delivery related queries’ between two matched groups across eight SDoH factors. 365
B.24 Differences in percentage points for changes in ‘Food delivery related queries’ between two matched groups across eight census variables. 366
B.25 Percent change in ‘Food assistance related queries’ between two matched groups across eight SDoH factors. 367
B.26 Differences in percentage points for changes in ‘Food assistance related queries’ between two matched groups across eight SDoH factors. 368

LIST OF TABLES

Table Number Page

3.1 Descriptive statistics of Bing data for human needs analysis 28
3.2 Top 12 need subcategories with the largest increase 37
3.3 Top 12 need subcategories with the largest decrease 37

4.1 Site characteristics and interview participant roles 86
4.2 Challenges at the intersections between psychosocial and cancer care journeys . . 101
4.3 SCOPE system capabilities 131

5.1 Descriptive statistics of variables used in engagement modeling 214
5.2 *Engaged* model parameters for all system-initiated nudges 223
5.3 *Distract Chosen*, *Calm Chosen*, *Address Chosen* model parameters 226
5.4 *Engaged* model parameters for all intervention choices 227
5.5 *Liked* model parameters for engaged interventions 229
5.6 *Improved* model parameters for engaged interventions 230

A.1 Distribution of demographic variables for US ZIP codes in our dataset 320
A.2 Pearson correlation between Bing client rate and demographic variables for each ZIP code 320
A.3 Pearson correlation between Bing and Google search trends 321
A.4 Human need categories and subcategories with examples 328

B.1 Means and medians of eight SDoH factors across US ZIP codes in our dataset . . . 335
B.2 Split boundaries used to generate comparison groups 335
B.3 Pearson correlation coefficients between SDoH factors 335
B.4 Seven search categories examined and their example search strings and regular expressions 336
B.5 Summary of pre-matching balance assessment and sample sizes across SDoH factors 337
B.6 Summary of post-matching balance assessment and sample sizes across SDoH factors 337

B.7 Balance assessment between unmatched high (Treated) and low (Control) '% Black or African American Alone' groups . . . 338
B.8 Balance assessment between matched high (Treated) and low (Control) '% Black or African American Alone' groups . . . 338
B.9 Balance assessment between unmatched high (Treated) and low (Control) '% Hispanic or Latino' groups . . . 338
B.10 Balance assessment between matched high (Treated) and low (Control) '% Hispanic or Latino' groups . . . 339
B.11 Balance assessment between unmatched high (Control) and low (Treated) 'Median Household Income' groups . . . 339
B.12 Balance assessment between matched high (Control) and low (Treated) 'Median Household Income' groups . . . 339
B.13 Balance assessment between unmatched high (Control) and low (Treated) '% with Bachelor or Higher Degree Attained' groups . . . 340
B.14 Balance assessment between matched high (Control) and low (Treated) '% with Bachelor or Higher Degree Attained' groups . . . 340
B.15 Balance assessment between unmatched high (Treated) and low (Control) 'Population Density (per sq mile)' groups . . . 340
B.16 Balance assessment between matched high (Treated) and low (Control) 'Population Density (per sq mile)' groups . . . 341
B.17 Balance assessment between unmatched high (Treated) and low (Control) '% Unemployed' groups . . . 341
B.18 Balance assessment between matched high (Treated) and low (Control) '% Unemployed' groups . . . 341
B.19 Balance assessment between unmatched high (Control) and low (Treated) '% with Health Insurance Coverage' groups . . . 342
B.20 Balance assessment between matched high (Control) and low (Treated) '% with Health Insurance Coverage' groups . . . 342
B.21 Balance assessment between unmatched high (Control) and low (Treated) '% with Internet Access' groups . . . 342
B.22 Balance assessment between matched high (Control) and low (Treated) '% with Internet Access' groups . . . 343

ACKNOWLEDGMENTS

It feels surreal to be writing the acknowledgment section of my dissertation. A tsunami of emotions and gratitude washes over me as I reflect on so many people that have touched my life in profound ways.

I want to begin by thanking my husband, Jeremy Im, who has been an unwavering foundation for me since the day we met. He has supported my ambitions and quest for knowledge and a career without question. I am eternally grateful for his loving support. Of course, I owe him countless laundry, dishwashing, and child chauffeuring that accumulated over the years. Knowing that he will gladly pick up my slack has given me stability in life that I cannot thank him enough for.

Looking back at my PhD years, I am perhaps more proud of the fine young woman that my daughter has become than the dissertation work itself. I am grateful for Leah's very existence as she continues to inspire me to make a difference in the world. In so many ways, she has been my "reverse parent" in encouraging me to get my PhD done, creating space for me to focus, and putting aside her own needs for mine.

I would like to thank my mom, Sun Suh, from whom I have fortunately or unfortunately inherited the constantly multitasking and overachieving nature. She not only kept up with her job and the community she serves but also took care of my family, my pets, my yard, and my house while I was pursuing my endeavors. She teaches me every day what it means to be a selfless parent. I am forever grateful to my dad, Chong Suh, who has sacrificed everything to bring me and my sister to the US and to allow me to shine.

I am fortunate to have five distinguished members on my dissertation committee whose dedication to their respective research areas shaped my research. I am especially grateful for their support during my PhD. Armed with their advice and insights, I exit my PhD as a proud researcher ready to serve the community and society.

Mary Czerwinski helped fan my HCI flame. I met Mary when I was just a master's student trying to grasp what it means to apply HCI theories into practice. Since then, Mary has been my biggest role model and advocate. She encouraged me to get a PhD, supported me through thick and thin, and gave me ample space to grow and give back. Mary welcomed me into VIBE, now the HUE group, and taught me how to be empathetic through technological innovation and human-centered design. She has given me the spotlight to shine and promoted me when I was afraid to face the world. With Mary by my side, I feel like I can take on the world. I thank Mary for giving me the space to pursue my PhD while working.

I thank James Fogarty and Dan Weld for welcoming me into the UW CSE department as a PhD student. As a naive prospective student, I was so sure to become a human-AI collaboration researcher. I had no clue at the time that I would turn my attention toward the mental health space. I am grateful to James for being my advisor and guiding me in this discovery process toward the SCOPE project. In research conceptualization, interdisciplinary collaborations, software engineering, and even parenting, I appreciate James's keen ability to step back and see the bigger picture as an architect. James has taught me how to navigate the ambiguities and challenges of interdisciplinary collaboration. Without James, I would not be a part of such an important project as SCOPE.

I appreciate Tim Althoff for taking me on as one of his first PhD students. Tim truly embodies his research passion for empowering people. If we were to analyze our advising meeting conversations through NLP, I am certain that Tim would score high on all factors indicative of empathic coaching. Tim has put me on a pedestal on so many occasions even when I felt I didn't deserve it. I am grateful that Tim has helped carve out a collaborative project between Microsoft and UW with Eric Horvitz and Ryen White. Tim has also demonstrated that perseverance with reviewers pays off in the end. Despite my imposter syndrome, Tim has shown me that my data science skills (which he helped hone) can add so much value to society.

Eric Horvitz has mentored me since I joined Microsoft Research as a developer and guided me in amplifying the voices of non-researchers and ultimately in developing my research agenda. Despite his busy schedule, Eric has always made time to share his ideas, give advice, and promote my work. I appreciate Eric's honest reflections on his years of research experience that I can tap into as I strategize on the next phases of my career.

Patricia Areán inspired me that an HCI researcher can make a real difference in the clinical mental health domain. My formative years as a PhD student were shaped by the vision of the Alacrity center that Pat founded, and I hope to continue being a steward of that vision.

I would like to walk down memory lane to thank those that helped guide me to where I am today.

I would not have been introduced to the technology sector without Rand Myles, who saw something in a Physics PhD student and decided to hire me into Microsoft. I would not have been introduced to Microsoft Research without Curtis Wong, who also saw something in a Physics PhD student and brought me to the wonderful world of research. I am indebted to Rand Myles, Arnold Blinn, and Zhangwei Xu who have helped me discover and cultivated my love for software development. They instilled in me a passion for technology as well as an empathy for people who use the technology, and I am thankful for their guidance and mentorship throughout the years.

I would not have been re-introduced to Microsoft Research without AJ Brush, who suggested that I pursue a career as a research developer, and Greg Smith, Gavin Jancke, and Patrice Simard convinced me that I could make a difference in research. I am grateful for Patrice who took a chance to hire me into the CHIL group with zero experience in HCI or ML and gave me full freedom to explore the research world. I appreciate Patrice for continuously reminding me to "write a dissertation, not *the* dissertation." I thank Saleema Amershi, who introduced me to HCI research and showed me how exciting human subjects research can be, and Carlos Garcia Jurado Suarez, who encouraged me to drive my research projects. I am grateful for Jerry Zhu who helped me run my first ML experiment that led to a publication. I learned so much ML from Max Chickering and Chris Meek and software engineering from John Wernsing, Mo Wang, Soroush Ghorashi, and John Verwey, and those learnings shape me to this day. Alicia Edelman Pelton taught me how to navigate Microsoft Research. I thank Steven Drucker and Bongshin Lee who showed me the power of visualization.

I thank the HUE group for giving me a second home and taking a chance with me. Rarely do you find a group of people who are so aligned with you and who go above and beyond their responsibilities to support you. It is not an exaggeration to say that the HUE group participated in my PhD with me and that they deserve this PhD as much as I do. I thank Daniel McDuff and Javier Hernandez for introducing me to affective computing and helping me discover what sensing technology is capable of. With his default response, "most certainly," Javier has taught me what it means to be a true team player. He also continuously reminded me that I need to prioritize finishing my PhD, which I am super thankful for. I thank Kael Rowan for his incredible development skills but also for his honesty and courage to speak up when we are straying from the path. I thank Daway Chou-Ren and Judith Amores for opening my eyes to a world that I did not know was possible. I would like to thank Gonzalo Ramos for being on my "nope" board. Since the day he joined Microsoft Research, Gonzo has poured hours of his time (and chocolates) into mentoring and supporting me through my hardships. I thank Gonzo for his unlimited supply of chocolates that kept me going, his openness to being my sounding board whenever I asked, and his courage to challenge anything that does not seem right.

I would like to thank Ryen White who was integral to the success of my Bing-related work. Ryen has opened the possibility of using Bing data for good and supported me in the execution of my project even through the painful review processes. I cannot thank Ryen enough for his support behind the scenes for my career development as well.

I am also grateful for friends and collaborators at Microsoft that have encouraged me to keep going in my PhD. I am fortunate enough to receive mentorship and guidance from Megan Saunders, Desney Tan, and Rico Malvar. Shamsi Iqbal, Koustuv Saha, Jenna Butler, Ali Benter, and Kathy

Dixon have shown me the real joy in collaborating on well-being research. I want to thank my collaborators in the central engineering group who helped me with my dissertation research: Marah Abdin, Tracy Tran, ThuVan Pham, and Vanessa Rodrigues. I thank the women in Microsoft Research who hugged and cried with me to keep going. Piali Choudhury has been an invaluable friend that supported me in my transition from a developer to a researcher. Robin Moeur has been a fierce advocate for me, and I admire her courage and endless supply of advice. I thank Ebele Okoli, Megan Lawrence, Wendy Chisholm, and Ioana Tanase in AI for Accessibility group who gave me the opportunity to expand my horizons.

I am fortunate to have received support from both Microsoft and UW, and there are many members of UW who have been instrumental in my PhD journey. I am grateful to be a member of two awesome research groups at UW, fogies and bdata, that I could tap into for advice and mentorship when my research projects were especially unclear. I would like to thank Ravi Karkar and Jessica Schroeder for guiding me through the PhD process with brutal honesty and insights. Ravi has helped me remain incredibly calm with his pessimistic yet hopeful outlook. To this day, I am still in search of a research collaborator as accountable and organized as Jessie had been for me. I thank Elise deGoede Dorough for helping me hack away at the PhD process to make it work with my odd situation.

My collaborators in the SCOPE project have given me a once-in-a-lifetime opportunity to help patients. I thank Gary Hsieh for his guidance and feedback during my early project days and for giving me the space to really dive into the multistakeholder problem, and I thank Spencer Williams for being my partner in crime. I thank Jesse Fann for the opportunity to help patients with cancer and depression, and Amy Bauer and Ty Lostutter for their clinical guidance and insights. I am grateful to have met an amazing group of PhD students (and now faculty), Ravi Karkar, Anant Mittal, and Tae Jones. I thank Ben Zheng and Lydia Andris, without whom, the SCOPE project would not move forward.

I want to thank all of the interns who have given me “reverse mentorship” through their enthusiasm for research and their unique perspectives and expertise. I thank Ismini Lourentzou, Alexandra Schofield, Joyce Passananti, Anna Wang, Nan-Chen Chen, Felicia Ng, Ted Grover, Napol Rachatasumrit, Samiha Samrose, Esther Howe, Mehrab Bin Morshed, Haekyu Park, Subigya Nepal, Anna Kawakami, Talie Massachi, Matthew Joerke, Yasaman Sefidgar, Vedant Das Swain, Harish Haresamudram, Sachin Pendse, and Robert Lewis.

This research could not have been possible without funding and support from the National Institutes of Health under awards P50MH115837, R01LM012810 and R01MH125179, the National Science Foundation under awards IIS-1813675, IIS-1901386, NSF CAREER IIS-2142794, and CNS-2025022, Bill & Melinda Gates Foundation under award INV-004841, the US Office of Naval Research under award #N00014-21-1-2154, the Allen Institute for Artificial Intelligence, the University of Washington Garvey Institute for Brain Health Solutions, and Microsoft Research.

DEDICATION

To Jeremy and Leah,

I am infinitely grateful for your love and for inspiring me to never give up.

Chapter 1

INTRODUCTION

Mental health disorders are a leading cause of disability and death worldwide, with approximately one in five of US adults experiencing a mental illness in a given year [298]. Despite high prevalence and negative impact, mental health disorders are severely under-treated [222]. Less than 25% of adults with a diagnosable mental health disorder receive professional psychotherapy [222], and 20% to 50% of patients are known to prematurely drop out of therapy [405, 374, 375]. At the patient level, significant barriers to treatment include high financial burden, lack of insurance coverage, lack of access to quality care [330, 202, 125], stigma, and competing life demands all contributing to not seeking or discontinuing care. Lack of professional training, insufficient time and resources, competing demands of comorbidities, and inadequate reimbursement systems contribute to added challenges that healthcare providers face [222]. Mental health disorders are also associated with increased risk for chronic diseases due to unhealthy behaviors such as lack of physical activity, poor diet, or substance abuse [319]. Furthermore, people with co-morbid physical and mental disorders encounter additional barriers (e.g., disease management, treatment complexity, reduced quality of life) that limit their ability to care for their mental health, which then adversely affects adherence to treatment [37]. Unfortunately, those that are overwhelmed with a multitude of physiological and economic burdens are especially unable to prioritize seeking support where symptoms of stress, anxiety, and depression are often overlooked and in constant tension with life demands and disruptions.

Even though many evidence-based psychosocial interventions (EBPIs) have been proven to be effective in treating mental health conditions, they suffer from a “research-practice gap,” leading to low adoption, low engagement, and low fidelity in adhering to the original guidebook [252]. One key contributor to such a gap is that the EBPIs are often evaluated in controlled settings when, in reality, these interventions are practiced in non-traditional settings (e.g., work, schools). EBPIs designed in controlled settings without consideration of unanticipated contextual challenges make it difficult for patients to engage in interventions in their unique situations and, therefore, have the potential to push people towards deprioritization of mental health care. Although evidence-based practice in psychology guidelines recommend “the integration of the (1) best available research with (2) clinical expertise in the context of (3) patient characteristics, culture, and preferences,” little progress has been made in incorporating patient values, culture, and preferences in designing and adapting interventions and treatment in comparison to the progress made in empirical research and clinical expertise [392]. To design contextually fit interventions and practice, it is critical to examine the “competing demands” that an individual experiences [301] that shape the access and engagement of those interventions while targeting both treatment outcomes and engagement and usability metrics. The shift in perspective from disease-specific to person-centered care requires cultural and contextual adaptations to existing EBPIs [26, 394] as well as interventions that exhibit dynamic and adaptive properties for supporting flexible decision-making in any given context [3, 256].

In recent years, technology has played a pivotal role in improving engagement in mental health care. Digital behavioral health technologies (e.g., telepsychiatry, text-based counseling, mHealth apps) create new opportunities to expand the reach of evidence-based practices beyond traditional clinical settings [144, 149, 151, 250] while also providing new opportunities for contextual understanding. The collection of multifaceted and high-fidelity digital usage and behavioral data can help characterize real-world contexts, observe physiological and behavioral patterns, and measure the effectiveness of mental health solutions in vivo, which can be used to ensure that people receive the best possible support in their situated contexts [292, 291, 355, 386]. Despite technological innovations and promises to address mental health challenges, such interventions ultimately need to consider the context within which they are deployed because access to technology itself may be a barrier to mental health care among the multitude of competing needs that discourage engagement in mental health and well-being care. Therefore, research in technology-facilitated mental health care and engagement must also shift towards a holistic approach to intervention design and delivery informed by contextual understanding.

1.1 THESIS STATEMENT

In my research, I seek (1) to explore the nature of competing needs and resource tensions in situated contexts that influence people's engagement in mental health and well-being activities, (2) to design and build contextually appropriate interventions that simultaneously promote engagement in and efficacy of mental health and well-being care, and (3) to deploy and evaluate these interventions in real-world contexts for dynamic and continuous improvements to interventions and their delivery.

My dissertation research, therefore, demonstrates the following thesis statement:

> *Engagement in mental health and well-being care can be improved through contextual and continuous adaptation of interventions by using human-centered and computational methods to understand stakeholder needs at personal and population scales.*

I claim that using *human-centered and computational methods* can reveal the tensions within the multi-level system of *stakeholder and well-being needs* that get in the way of engaging in mental health and well-being activities. I also claim that such contextual understanding, at both *personal and population scales*, can support the appropriate *design and adaptation* of technology-facilitated interventions aimed at *improving engagement in mental health and well-being care*.

1.2 DISSERTATION OVERVIEW

My dissertation has spanned three key research activities across three mental health and well-being contexts in support of my thesis.

First, the demonstration of my thesis follows three key research activities that are typically seen in the human-centered design process [176, 278].

1. **Framing and understanding:** Within the framing and understanding activity, the goal is to use holistic frameworks to understand the conflicts within the multi-level system of human

needs at personal and population scales. Here, I leverage human needs framework [264, 269], social determinants of health [419], and parallel journeys framework [400] to gain an understanding of needs that compete with the prioritization of well-being activities.

2. **Designing and adapting:** Within the designing and adapting activity, the goal is to adapt and digitize core components of evidence-based interventions for improved usability and engagement. Here I digitize Behavioral Activation [208], Collaborative Care Model [17], Cognitive Behavioral Therapy [73], and Dialectical Behavioral Therapy [251] into contextually fit solutions that accommodate the conflicts identified in the prior research activity. I identify modification targets for these interventions that can be enhanced through the use of technology.
3. **Deploying and evaluating:** Within the deploying and evaluating activity, the goal is to conduct longitudinal deployments and identify opportunities for continuous adaptation and personalization of interventions. Here, I deploy the systems that I build for user studies and randomized clinical trials (RCTs) and analyze the data from the deployments to propose design considerations for both technical and non-technical systems.

Second, I apply this process across three mental health and well-being contexts – (1) the COVID-19 pandemic, (2) co-morbid cancer and depression, (3) workplace stress – to demonstrate my thesis across the scale of analyses from momentary to longitudinal, from individuals to populations, and from qualitative to quantitative.

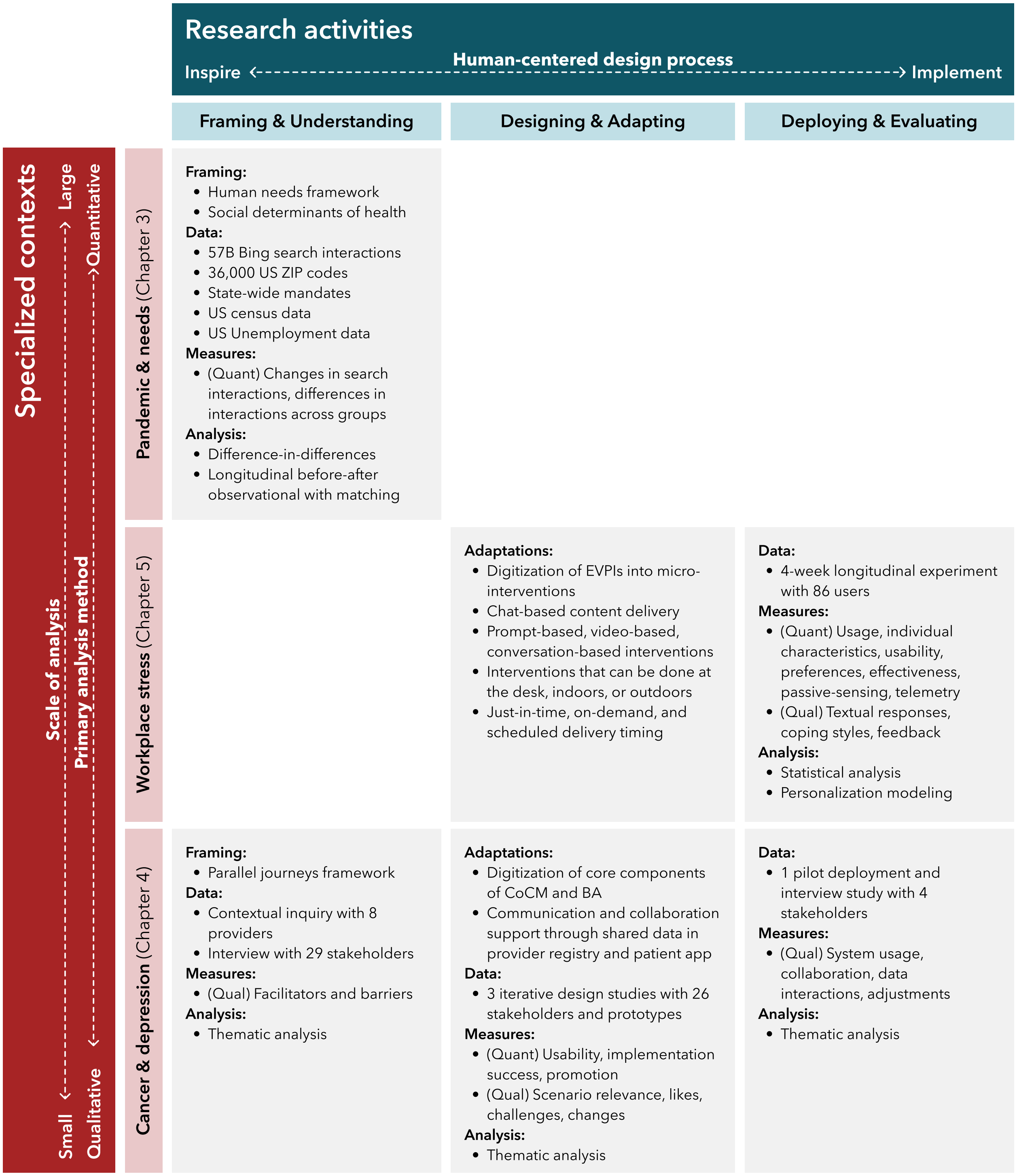
1. **COVID-19 pandemic:** The COVID-19 pandemic has disrupted the lives of people across the world in both digital and offline realms. Here, I leverage computational methods to understand population-scale shifts in online search behaviors during the pandemic. My analysis reveals a general shift towards basic needs away from higher-level aspirational needs as well as a disproportionate change in a community's use of digital resources across several socioeconomic and environmental offline factors.
2. **Co-morbid cancer and depression:** Depression is highly prevalent but severely under-treated among patients with cancer. Here, I leverage qualitative methods to understand multistakeholder challenges that arise at the intersection of cancer and psychosocial care

journeys. To address those challenges through technology support, I digitize and adapt core components of evidence-based care strategies aimed at improving engagement in depression care. I then deploy the system for an evaluation in a series of user studies and a multi-year RCT.

3. **Workplace stress:** Workplace stress is prevalent and costly, but integrating stress management strategies directly into the workday can be challenging. Here, I digitize evidence-based psychosocial interventions into bite-sized, digital micro-interventions that can be delivered in the moment through the use of passive sensing technologies. I then deploy the system to evaluate how momentary contextual information can influence engagement in interventions at work.

My dissertation, therefore, makes several types of contributions. I make theoretical contributions through the development of the parallel journeys framework as a design framework for examining and characterizing complex co-morbidities and care contexts. I make methodological contributions through the development of the human needs computation framework as a way to capture changes in digital engagement that span a broad spectrum of human needs and through the development of the longitudinal before-after observational method that quantifies disproportional changes in digital engagement. I make artifact contributions by designing and developing a collaborative behavioral activation system and a just-in-time micro-intervention system. I make several empirical – both qualitative and quantitative – contributions. I quantify the changes in human needs during the pandemic and observe the second-level digital divide at population scales. I characterize the challenges that multiple stakeholders face during co-morbid cancer and depression care and during the deployment of a collaborative behavioral activation system. I measure the efficacy of digital stress-reduction micro-interventions at work and identify factors that influence engagement in the moment. I leverage my empirical findings to contribute design implications for public health policies, the implementation of technology-enhanced collaborative care, and a just-in-time adaptive interventions (JITAI) aimed at improving engagement in mental health and well-being care.

Figure 1.1 summarizes the completed studies included in my dissertation.



Research activities

Human-centered design process

Inspire <--> Implement

Specialized contexts

Scale of analysis: Small <---------------------------------> Large

Primary analysis method: Qualitative <---------------------------------> Quantitative

	Framing & Understanding	Designing & Adapting	Deploying & Evaluating
Pandemic & needs (Chapter 3)	**Framing:** • Human needs framework • Social determinants of health **Data:** • 57B Bing search interactions • 36,000 US ZIP codes • State-wide mandates • US census data • US Unemployment data **Measures:** • (Quant) Changes in search interactions, differences in interactions across groups **Analysis:** • Difference-in-differences • Longitudinal before-after observational with matching		
Workplace stress (Chapter 5)		**Adaptations:** • Digitization of EVPIs into micro-interventions • Chat-based content delivery • Prompt-based, video-based, conversation-based interventions • Interventions that can be done at the desk, indoors, or outdoors • Just-in-time, on-demand, and scheduled delivery timing	**Data:** • 4-week longitudinal experiment with 86 users **Measures:** • (Quant) Usage, individual characteristics, usability, preferences, effectiveness, passive-sensing, telemetry • (Qual) Textual responses, coping styles, feedback **Analysis:** • Statistical analysis • Personalization modeling
Cancer & depression (Chapter 4)	**Framing:** • Parallel journeys framework **Data:** • Contextual inquiry with 8 providers • Interview with 29 stakeholders **Measures:** • (Qual) Facilitators and barriers **Analysis:** • Thematic analysis	**Adaptations:** • Digitization of core components of CoCM and BA • Communication and collaboration support through shared data in provider registry and patient app **Data:** • 3 iterative design studies with 26 stakeholders and prototypes **Measures:** • (Quant) Usability, implementation success, promotion • (Qual) Scenario relevance, likes, challenges, changes **Analysis:** • Thematic analysis	**Data:** • 1 pilot deployment and interview study with 4 stakeholders **Measures:** • (Qual) System usage, collaboration, data interactions, adjustments **Analysis:** • Thematic analysis

Figure 1.1: Scope of research work in my dissertation including data and methods used in each.

1.3 ORGANIZATION OF THE DISSERTATION

My dissertation has focused on understanding three mental health and well-being contexts through human-centered and computational methods, designing technology-facilitated and contextually fit interventions, and deploying those interventions to evaluate their efficacy in those contexts. The overall organization of my dissertation is as follows.

In Chapter 2, I provide prior work in mental health and well-being interventions and the relevant technology contributions to improving engagement. I also provide background information on three specialized well-being contexts that I examine.

Chapter 3 describes my work in the context of the COVID-19 pandemic and online information-seeking behaviors. I first illustrate how understanding the context at population scales can highlight areas of competing needs. Then, I investigate how digital engagement in resources across health, educational, economic, and social needs were not experienced equally by populations defined by various socioeconomic and environmental factors. This work suggests that population-scale monitoring of online information-seeking behavior can highlight specific at-risk populations to inform policies and recovery efforts for current and anticipated needs. Section 3.1 was originally published at the 14th ACM International Conference on Web Search and Data Mining [402], and Section 3.2 was originally published in Nature Communications [403].

In Chapter 4, I present my work in the clinical context of patients with co-morbid cancer and depression. I propose the parallel journeys framework to characterize the challenges that arise at the intersection of cancer and psychosocial care journeys and develop a technology-facilitated collaborative behavioral activation system, called SCOPE. By evaluating the system in a series of user studies and a clinical pilot study, I highlight technical and non-technical design implications for integrating collaborative systems in complex clinical care settings. Section 4.1 was originally published in the Proceedings of the ACM on Human-computer Interaction (CSCW1).

In Chapter 5, I describe my work in the workplace context to develop stress-reduction micro-interventions and a just-in-time system that delivers them at work. By deploying the system in a four-week user study, I demonstrate that an adaptive workplace stress intervention system can lead to workplace stress reduction. I then identify contextual and individual factors that influence engagement and improvement with implications for future JITAI systems. Section 5.1 was originally published at the 2022 CHI Conference on Human Factors in Computing Systems, and the contents of Section 5.2 currently in submission.

Finally, Chapter 6 summarizes how the presented research work provides evidence for my thesis statement. I also discuss future challenges and opportunities for investigating and improving engagement in mental health and well-being care that spans individual, collaborative, and public health contexts.

Chapter 2

RELATED WORK AND BACKGROUND

My dissertation research explores the challenges and barriers to engaging in mental health and well-being care to inform the design of technology enhancements and adaptations to evidence-based interventions aimed at improving engagement. In this chapter, I provide an overview of mental health and well-being interventions and the need for adaptations (Section 2.1), followed by the relevant prior work in the design of technology support for mental health and well-being (Section 2.2). I then describe three well-being contexts in which I examine the tensions among a multitude of human needs and design technology support for mental health and well-being management (Section 2.3).

2.1 MENTAL HEALTH AND WELL-BEING INTERVENTIONS

Mental health and well-being interventions and strategies have been developed and validated across various domains and contexts. However, mental health treatments are still severely underutilized. Less than 25% of adults with a diagnosable mental health disorder receive professional psychotherapy [222], and 20 to 50% of patients in therapy prematurely drop out of therapy [299, 375, 374, 405]. Barriers for patients in receiving treatments include lack of access to quality care, financial burdens due to lack of insurance coverage, clinician shortage, stigma, and perceived ineffectiveness [125, 202, 287, 330]. Behavioral health providers (BHPs) face various challenges including the lack of professional training, insufficient time and resources, competing demands of comorbidities, and inadequate reimbursement systemse [222]. Even for patients who receive care, unfortunately, they struggle to remain engaged in care or receive consistent care with 20 to 50% of patients known to prematurely drop out of therapy [405, 374, 375, 299].

Despite prolific research and development of evidence-based interventions and strategies, translating research in evidence-based psychosocial interventions (EBPIs) into clinical practice or everyday life has been met with several challenges. When the interventions are developed and evaluated in highly controlled settings (e.g., a sequence of consistent, in-person therapy sessions), implementing these programs in real-world contexts has often failed to replicate the "in-the-lab" effectiveness [256]. In addition, these interventions are evaluated with a singular focus on disease-centric target outcomes with little or no regard to their usability or how people engage with them "in the wild." In reality, many of these interventions are practiced in non-traditional settings such as work, schools, primary care, or community service settings [253].

Integration of EBPIs into everyday lives outside of therapy contexts also relies solely on user motivation without support for compliance and adherence [70]. Due to their poor contextual fit, these EBPIs often suffer from a "research-practice gap," leading to low adoption and low fidelity in adhering to the original guidebook [252]. Furthermore, little progress has been made in incorporating patient values, culture, and preferences in designing and adapting interventions and treatment in comparison to the progress made in the empirical research and clinical expertise [392]. Attempts at integrating EBPIs without the consideration of contextual challenges would continue to push people towards deprioritization and under-utilization of mental health care. Therefore, improving the usability of and engagement with EBPIs should be one of the top target metrics of research and development of mental health and well-being interventions [25, 120, 128, 252, 267, 282].

Prior research in the implementation science field has examined ways to modify interventions in reaction to unanticipated challenges in practice [27]. When mental health and well-being interventions and strategies are met with unanticipated challenges that interfere with their effectiveness, they can be "modified" in reaction to such challenges [27]. These interventions could also be "adapted" deliberately and proactively ahead of deployment [394, 393]. For example, cultural adaptations consider the compatibility of language, cultural practices, meanings, and values in its modifications [26]. Contextual adaptations consider the culture as well as the context of use (e.g., accessibility, acceptability, structural barriers [70] or child development stages [256]) to improve the contextual relevance of interventions [70]. While cultural adaptations commonly apply to the intervention content (e.g., language), contextual adaptation encompasses how the interventions are used with usability, engagement, and retention as metrics to improve the contextual relevance of interventions [70]. Furthermore, interventions that are more adaptable have a higher chance of success in meeting the needs stemming from shifting contexts and situations [256]. Interventions, therefore, need to exhibit dynamic and adaptive properties, such as modular designs and identification of core components, to support flexible decision-making in a given context [3, 256].

My research examines critical gaps in bringing evidence-based mental health and well-being interventions into real-world, situated contexts using human-centered and computational methods. I first characterize the contextual challenges and barriers to engaging in mental health and well-being activities. Informed by these challenges, I then identify core components of interventions and design targeted adaptations aimed at improving flexibility and ease of use in the given contexts. In these situated contexts, I also seek technology design opportunities to enhance intervention delivery and consumption. Therefore, my dissertation research contributes to the growing literature on digital and contextual adaptations of interventions.

2.2 TECHNOLOGY SUPPORT FOR MENTAL HEALTH AND WELL-BEING

Researchers have investigated technological support of mental health from a variety of perspectives. Early calls for research at the intersection of mental health care and human-computer interaction (HCI) [101] have been met with over a decade of research investigating technological support for diverse clinical needs [352], such as anxiety disorders [102, 438], bipolar disorder [25], ADHD [235], phobias [456], schizophrenia [434], and depression [283, 363, 389]. In the context of mental health and well-being, innovators, researchers and therapists have explored the potential for new technology to overcome key barriers to mental health care access and to improve overall mental health and well-being outcomes for all. For example, computer-assisted therapy methods have been examined in a variety of contexts, including for anxiety disorders [102, 438], bipolar disorder [25], ADHD [235], phobias [456], schizophrenia [434], and depression [389, 283, 363]. The use of technology to help support mental health treatment has significant potential and has proven highly effective in a number of contexts [352].

In particular, technology-delivered interventions have garnered attention for their ability to offer constantly-available content that people can engage with [72, 335, 336] to reduce mental health symptoms [284, 363]. Given the potential for significant impact, many innovators and researchers have created and studied the efficacy of smartphone-delivered mental health interventions via RCTs [225, 244] and a myriad of consumer-oriented mental health applications exist to provide a range of support online, such as self-guided meditation or symptom management. However, most available mental health apps are not evidence-based [191, 294], and those that have been rigorously studied – although shown to be effective in improving relevant outcomes [225, 244] – also have high attrition and low adherence [134, 243].

Integrating technology into mental health and well-being solutions has introduced new ways to adapt existing EBPIs. During the digital translation of EBPIs, technology-delivered interventions can be broken into specific modules or skills [363]. They can also be adapted to be narrow in scope and short in length (e.g., 1-minute meditation). These "digital micro-interventions" leverage technology affordances to provide individual components of traditional psychotherapy focused on managing proximal symptoms (e.g., relaxation for stress) in the hopes of achieving broad, distal objectives (e.g., overcoming depression) [35]. This modularization of EBPIs makes it easier for individuals to access and utilize interventions whenever they want especially between therapy sessions.

Furthermore, the modularization of EBPIs can enable dynamic adaptations, especially when coupled with computational approaches for behavioral, symptom, or contextual understanding. For example, systems delivering these modular interventions can take advantage of the usage and interaction data for personalization, improving its recommendations for activities that are likely to be effective [364], used [218], preferred [305], or performed at the right time [292, 354, 381]. The fullest extension of such personalization, termed just-in-time adaptive interventions (JITAI), has been introduced to continuously adapt and deliver personalized, contextualized, and adaptable interventions incorporating dynamic human behavior data captured through ubiquitous sensing technologies [292, 355, 386, 47]. This concept has been pursued in other health contexts (e.g., promoting physical activity [169], stress management [201], and weight management [387]), with the recent interest in applying them to positive coping skill use [206]. JITAI approach holds great potential for improving adherence and engagement, wherein micro-intervention delivery timing, as well as the content, can be continuously adapted based on ecological momentary assessment (EMA) or other passively sensed data [355, 381].

Technology support for mental health and well-being care expands opportunities to leverage data to improve the efficacy of interventions. For example, prior work has examined evaluating counseling sessions through natural language processing and machine learning approaches to differentiate high-quality and low-quality counselors or counseling sessions [11, 312, 406]. Research has investigated monitoring and predicting mental health symptoms through mobile phones and ubiquitous sensors [40, 433, 284] and social media posts [106, 105, 18, 257, 461, 350]. A large body of work has examined how machine learning techniques can support the detection, diagnosis, and treatment of a myriad of mental health conditions [369, 409]. Researchers have also leveraged usage data to characterize user behaviors [83] and investigated personalizing mental health interventions by recommending activities to manage stress [354] or prevent negative moods [189] based on an individual's past sleep, diet, and activity data; by recommending interventions based on an individual's personal characteristics and context [305].

Despite technological innovations to address mental health challenges and the potential for digital mental health solutions to extend into the fabric of our lives, they often fail to address individual, institutional, and infrastructural level challenges in mental health care, such as challenges for providers connecting patients to local resources, a lack of mental health resources for low-income patients, a legal system that does not address mental health needs, or a lack of reimbursement mechanisms [51, 285]. Such challenges remind us that digital interventions need to be designed with consideration for the context in which they are deployed and that even access to technology itself may be a barrier to mental health care. Therefore, my dissertation places the context of use at the center and examines the role of technology in supporting engagement in mental health care. In order to build adaptable systems that meet the dynamic and personalized needs of a given context and that optimize flexibility, usability, and engagement, my work leverages the modularity of the core components of EBPIs in their digital translation and the adaptability of intervention timing and delivery through ubiquitous sensing technologies.

2.3 MENTAL HEALTH AND WELL-BEING CONTEXTS

In my dissertation, I investigated three well-being contexts to characterize tensions in needs that get in the way of engaging in mental health and well-being activities. Challenges identified in each of these contexts inform the design of technology enhancements and contextual adaptations to interventions that promote engagement in mental health and well-being care. Here, I provide background information on these problem contexts and on prior methodological approaches to address those challenges in relation to my research.

2.3.1 The COVID-19 Pandemic

The COVID-19 global pandemic has presented challenges across physical health [440] as well as societal (e.g., disparities [88]), economic (e.g., unemployment [23, 91]), and psychosocial (e.g., stress, anxiety [314], loneliness [420]) realms. Because of the multi-faceted influence of the pandemic on health and well-being, recent work has called for understanding the multi-level system of humans needs for pandemic response strategies [347] and examining the unintended socioeconomic consequences of the pandemic [39].

Theories about basic human needs have been discussed for close to a century [318]. In particular, Maslow's hierarchy of needs [264, 265] has been applied in numerous domains [459, 424, 79, 289], despite criticisms of the validity of the theory [429, 59] and controversies surroundings its relationship to Blackfoot Native American tribe [38, 139, 64]. These theories are aimed at providing a holistic understanding of human needs. The characterizations of a broad spectrum of needs from these models are increasingly relevant during the pandemic [347].

Most studies on human needs use survey-based methods [79, 407, 266, 281]. Others have applied human needs theories in computational social science [9, 459, 247], but they focus deeply on specific topics (e.g., consumer behavior, well-being) and leverage publicly available social network data. In contrast, our work introduces a computational methodology to extract a full spectrum of human needs at population scales, which is critical for aligning policies with societal needs that they are intended to support.

In particular, we harness search interaction data from web search logs. Web search logs are routinely collected on a near real-time basis and at large scales, providing unique opportunities to examine digital behaviors across a wide range of topics, geographies, and subpopulations as well as highlighting potential barriers and changes to such engagement behaviors [127]. In fact, web search logs have enabled studies of human behaviors across many different domains [15, 152, 442, 437], times [12, 13, 309, 146], locations [443, 349], and to make inferences about the future or to identify risk factors [104, 86, 228, 304, 445]. In the context of the COVID-19 pandemic, such data has stimulated a prolific range of research on physical [242, 228, 153], psychological [414, 65], and socioeconomic [163, 4] well-being [402].

Socioeconomic and environmental factors play a significant role in the health and well-being of individuals and communities [140, 455, 450]. Despite pandemic-driven efforts to close the long-term and emergent health equity gap [455], studies during the COVID-19 pandemic have also demonstrated that socioeconomically and environmentally marginalized subpopulations have been disproportionately and negatively affected by the disease [81, 88, 458, 463]. Unfortunately, such disparities are also reflected in digital access and engagement [143]. Digital inequalities manifest in multiple levels as the differences in (1) the access to technology or the quality of access (i.e., first-level digital divide) [423], (2) the usage of digital technologies and skills (i.e., second-level digital divide) [172, 174], and (3) the ability to translate the use of digital technologies into favorable outcomes (i.e., third-level digital divide) [74, 338, 422]. The inability to digitally engage in more "capital-enhancing" activities [74, 173, 174] can lead to negative offline consequences [180].

Prior work on understanding digital disparities has relied on costly surveys, interviews, or self-reports [10, 164, 219] that require direct engagement with the study population in order to prompt a recounting of their past behaviors rather than passively observing their actual behaviors. Datasets from specific service providers (e.g., Wikipedia [175], Zearn.org [465, 464]), domains (e.g., telehealth [124], eHealth [356]) or geographic areas (e.g., Northern California [356]) do not capture digital behaviors across a broad spectrum of human needs and subpopulations and at fine geo-temporal granularities. Macroeconomic measures, such as unemployment claims, do not capture potentially unmet needs or access barriers (e.g., confusion around unemployment benefits [367, 46, 114]). Leveraging web search interactions, on the other hand, enables us to model people' search interests which are reflective of their underlying resource needs [408, 321, 447]. These interests could include accessing critical digital resources such as online educational sites in response to school closures, online food delivery information in response to restaurant closures, online social interactions in response to physical distancing and travel restrictions, or online unemployment and economic assistance in response to economic instability during the pandemic.

In the context of the COVID-19 pandemic, my work leverages the centrality of web search in everyday life to better understand changes in human needs during the pandemic, how these changes are experienced differently by different subpopulations to further characterize the impact of socioeconomic and environmental factors on engaging in digital well-being activities.

2.3.2 Cancer and Depression

Depression is one of the most common challenges in patients with cancer [221, 317] with devastating consequences, such as increased mortality risk [214], longer hospital stays [296], increased risk of completed suicide [181, 280, 348], increased healthcare costs [268], decreased adherence to treatment recommendations and medications [215], and decreased quality of life [317]. Common contributors to depression in cancer include a patient's psychological reaction to phases of the cancer journey (e.g., diagnosis, treatment, recovery or relapse, survivorship or end of life), social stressors (e.g., loss of a job, financial burden), and physical side effects from cancer treatment (e.g., nausea, fatigue, hair loss) [317, 380]. Recent meta-analyses have found the prevalence of depression in patients with cancer to be as high as 24% [221, 317], much higher than the 8% annual prevalence in the general population [435, 391]. Clinical depression in some patients may go undetected by oncology providers and staff who lack specialized training to diagnose mental health conditions [382, 307, 273].

Despite its high prevalence and negative impacts, depression in patients with cancer is highly under-treated with 73% not receiving potentially effective depression treatment [430]. Patients in rural areas are more likely to receive inadequate access to mental health services compared to urban residents, partly due to a limited supply of behavioral health providers (BHPs), lack of psychiatric clinicians, and lack of on-site mental health services [330, 202, 125]. In addition, patients with co-morbid conditions (i.e., cancer and depression) encounter many barriers that limit their ability to care for their health (e.g., physical limitations, high complexity in medication management, financial constraints, compounding of symptoms) [37].

Because of the potential direct neuropsychiatric effects of cancers and their treatment on depression [317], access to BHPs trained specifically in oncology contexts is crucial in the treatment of depression in patients with cancer. The Institute of Medicine recommends direct integration of psychosocial services in cancer settings [302] as a cost-effective way to address inadequate treatment [277, 317], and the psychosocial oncology care framework [254] provides guidelines for achieving such recommendations. The Collaborative Care Management (CoCM) model is an evidence-based strategy for implementing these recommendations [238], a proven and cost-effective standard of care for treating depression, improving quality of life, and increasing adherence to cancer treatments [21, 410, 78, 126, 197] consistent with accepted clinical practice guidelines [19].

Collaborative care is based on several core principles [425]: (1) it is a *team-based* collaborative approach to psychosocial care with a triad of providers (i.e., as described next), (2) it is a *population-based* care approach responsible for the outcomes of a defined population of patients (i.e., patients identified as depressed), (3) it is a *measurement-based* care approach that uses validated outcome measures to monitor patients and to guide treatment decisions, and (4) it is an *evidence-based* approach that uses scientifically proven interventions that are effective in treating depression [135, 17].

Collaborative care centers around the patient. A triad of providers, each with a clearly defined role, collaborate to improve the patient's psychosocial health outcomes. Figure 2.1, adapted from the University of Washington AIMS Center, illustrates the roles and relationships of collaborative care team members in the cancer context. An *oncologist* is responsible for overseeing the overall care of the patient. A *psychiatric consultant* works with the care manager and oncologist to aid

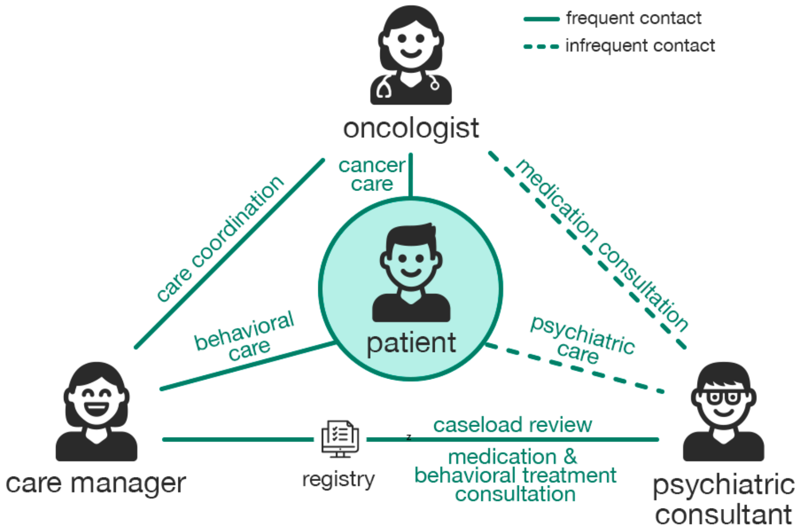


Figure 2.1: Illustration of Collaborative Care Model, which centers around the patient with a triad of providers (oncologist, care manager, consulting psychiatrist) collaborating to improve the patient's psychosocial health outcomes. The illustration is adapted from the University of Washington AIMS Center (http://aims.uw.edu/).

with assessment and diagnosis, to develop a biopsychosocial treatment plan, and to communicate recommendations for treatment (i.e., psychotropic medications and/or evidence-based psychosocial treatments) [32]. A *care manager*, who is typically a clinical social worker but can also be a nurse or psychologist, is embedded in the cancer center and consults with an on-site or remote psychiatrist with expertise in treating patients with cancer. They work closely with the care team (i.e., oncology, primary care, and other supportive care providers) to coordinate and deliver comprehensive care, provide patient education, deliver evidence-based psychosocial treatments, monitor patient progress using standardized scales, and provide systematic outreach to engage patients. [113, 168, 95]. Our work specifically examines depression care situated in the context of collaborative care.

Implementing collaborative care is also not without its own challenges. For example, recent reviews and analyses have highlighted: (1) low fidelity in adhering to the principles of effective collaborative care [425], (2) lack of clarity in role definitions and boundaries, (3) challenges in long-term sustainability, and (4) lack of standard care and communication pathways and tools [286, 370, 453]. Using the core principles of collaborative care as a basis for determining key tasks and processes, a comprehensive list of potential technology opportunities for enhancing collaborative care has been proposed [28, 326]. Prior studies have looked at augmenting the capabilities of collaborative care with technology through remote assessment [145, 224, 457], patient-provider communication [341], digital interventions [390], or telehealth [6]. However, many of these studies do not target the specific needs for depression care in cancer settings, and further research is needed to evaluate the effectiveness of such technologies on improving patient engagement in depression [325].

Using the core principles of collaborative care as a basis for determining key tasks and processes, a comprehensive list of potential technology opportunities for enhancing collaborative care has been proposed [28, 326], including a messaging platform for patient engagement, digital delivery of evidence-based psychosocial interventions and education resources, decision support tools with clinical pathways, remote self-assessment, and systematic monitoring and review of patients using a web-based registry and remote consultations. However, despite high interest in integrating technology into collaborative care and its promising benefits, further research is necessary in understanding the needs for and feasibility of such technologies in collaborative care for various stakeholders, especially for treating depression in cancer settings [325].

Because patients with co-morbid conditions (i.e., cancer and depression) encounter many barriers that limit their ability to care for their health (e.g., physical limitations, high complexity in medication management, financial constraints, compounding of symptoms) [37], co-morbid cancer and depression is a highly relevant domain to examine the tensions in needs and to identify opportunities to adapt existing interventions to fit the needs of the population. Given that collaborative care serves as the standard of care for patients with cancer and depression, my work aims to develop concrete technology design opportunities to support patient and care team experiences.

2.3.3 Workplace Stress

Over 83% of Americans suffer from work-related stress [293], which has a profound impact on well-being due to increased risk of mental and physical health disorders, decreased productivity due to absenteeism and burnout, decreased overall job satisfaction, and increased rates of stress-related accidents and employee medical, legal, and insurance costs [55, 62, 94, 444]. Many have theorized that workplace stress stems from the imbalance between job demands, control, or resources [110, 421]. This perceived imbalance not only highlights that the experience of work-related stress is highly personal and unique to individuals but also presents the conflict between various work and nonwork roles as one of the key contributing factors of stress. For example, when individuals face a conflict between work demands and personal needs (i.e., work-self conflict or work-nonwork conflict), they tend to give up their personal roles to eliminate the source of the conflict or over-commit to fulfill the responsibilities of both [111]. Such conflicts can have detrimental effects on physical and psychosocial health [111, 357]. In addition, the effects of work-nonwork conflict and related stress outcomes spill over into life outside of work, disrupting the overall well-being of workers [162, 357].

Intervention strategies for workplace stress are commonly grouped into three categories: primary, secondary, and tertiary [98, 372, 332, 333]. Primary strategies refer to action taken to directly change or eliminate stressors; these strategies often involve organizational-level changes (i.e., a culture shift), which might be difficult to achieve or even study because such changes are costly and may cause high-profile disruptions to the organization [98]. Secondary strategies are the most common; they target the individual experiencing stress and aim to detect and reduce their stress to prevent the development of chronic, stress-related mental and physical health issues. Individual talk therapy, workshops teaching stress reduction and time management, as well as mindfulness or meditation applications like HeadSpace are all popular examples of secondary strategies. Tertiary prevention concerns treatment for and recovery from stress-related mental and physical health issues through counseling or supportive services such as Employee Assistance Programs (EAPs). Some argue that addressing the source of stress through organizational changes, rather than assisting individuals, is necessary for long-term beneficial effects [98], and ample evidence points to the benefit of individual-focused interventions which are easier to implement and study [333]. Others suggest that such distinctions are not an important area for focus because individuals trained with stress management skills can bring about organization-level changes [60].

Of these strategies, individual-based stress management interventions (e.g., cognitive-behavioral skills training, meditation, exercise, etc.) have been shown to be effective on psychological, physiological, and organizational outcome measures, are easier to implement and study, and are perceived to be more effective [213, 333]. However, integrating individual-based stress management interventions into the workday is challenging due to work culture [123, 184], and putting stress management strategies into practice can be quite difficult psychologically, even in healthy workplace contexts that value employee well-being [179].

Recent efforts to integrate interventions into the workplace include technology-delivered mental health interventions [395]. Technology-delivered mental health interventions in the workplace have gained popularity because they are always available (unlike a therapist), and can be used without drawing attention to the user or disrupting a work environment. Much of the prior research on their effectiveness in work contexts mainly focuses on work-related outcomes (e.g., productivity, absenteeism) rather than mental health or well-being outcomes [395]. Many interventions evaluated at work are still based on self-guided training (e.g., 12-week web-based cognitive behavioral therapy) or apps (e.g., MoodHacker) that are not directly integrated into the work context. Instead, digital interventions that take the work context into account are still highly sought after [107]. However, most available mental health apps are not evidence-based [191, 294], and those that have been rigorously studied – although shown to be effective in improving relevant outcomes [225, 244] – also have high attrition and low adherence [243, 134].

Because of the unique experiences of individuals dealing with work-related stress and the work-nonwork interface, integrating technology-delivered interventions into the workplace requires a high degree of personalization. The system design also needs to be highly usable and effective in the context of work to avoid introducing additional sources of personal conflict and pressure in utilizing necessary self-oriented stress-reduction resources during work. My dissertation work targets stress reduction and engagement as outcomes and focuses on directly incorporating the work context into the design and delivery of digital interventions through the use of ubiquitous sensing technologies [275]. Because of the unique experiences of individuals dealing with work-related stress, I explore ways to integrate technology-delivered interventions into the workplace with a high degree of personalization. Ouside of my dissertation work, I also explore the impact of organization-level interventions (e.g., deployment of employee well-being sensing technologies) and how individual-based solutions can impact employee-employer relationships and team dynamics (e.g., sharing of well-being insights with managers or colleagues).

Chapter 3

UNDERSTANDING THE IMPACT OF THE PANDEMIC ON HUMAN NEEDS

The COVID-19 pandemic has disrupted the lives of people across the world. Despite pandemic-driven efforts to close the long-term and emergent health equity gap [455], studies during the COVID-19 pandemic have demonstrated that socioeconomically and environmentally disadvantaged subpopulations have been disproportionately and negatively affected by the disease [88, 463, 81], with threefold higher infection rates and sixfold higher death rates in predominantly Black US counties than in white counties [458]. Furthermore, challenges induced by the pandemic span societal (e.g., disparities [88]), economic (e.g., unemployment [23, 91]), and psychosocial (e.g., stress, anxiety [314], loneliness [420]) realms, making it critical to incorporate the multi-level system of human needs and well-being for pandemic response and recovery strategies [347]. During the COVID-19 pandemic, digital engagement in resources across health, educational, economic, and social needs also grew in importance because of lockdown mandates, social isolation, and economic burdens [402, 39, 20] as well as due to internet-based communication methods employed by public institutions, such as the online dissemination of COVID-related information by the World Health Organization [39]. Unfortunately, disparities in digital access also reflect socioeconomic and environmental dimensions of variation [143], and such inequalities may perpetuate existing disadvantages into a “digital vicious cycle” [39, 34]. Therefore, it is a public health priority to identify vulnerable populations who are severely and negatively impacted by the pandemic and to understand potential barriers to accessing critical resources [88].

In this chapter, I focus on online search behaviors during the pandemic as a context for which to *understand stakeholder needs*. I demonstrate that (1) *computational methods* can reveal these needs at *population scales*, and that (2) such understanding could inform the design of public health policies and interventions aimed at *improving engagement in online resources*. I conduct two studies that illustrate that harnessing the centrality of web search engines for online information access can reveal changes in human needs and disparities.

The first study [402] examines the role that the COVID-19 pandemic has played in shifting human needs (Section 3.1). In this study, we propose a computational methodology, inspired by Maslow's hierarchy of needs, that can capture a *holistic* view of relative changes in needs following the pandemic through a difference-in-differences approach that corrects for seasonality and volume variations. We apply this approach to characterize changes in human needs across physiological, socioeconomic, and psychological realms in the US, based on more than 35 billion search interactions spanning over 36,000 ZIP codes over a period of 14 months (Section 3.1.1, Section 3.1.2). The analyses reveal that the expression of basic human needs has increased exponentially while higher-level aspirations declined during the pandemic in comparison to the pre-pandemic period (Section 3.1.3, Section 3.1.4). In exploring the timing and variations in statewide policies, we find that the durations of shelter-in-place mandates have influenced social and emotional needs significantly (Section 3.1.5). The study also finds potential barriers to addressing critical needs, such as support for unemployment and domestic violence (Section 3.1.6). This approach and study results suggest that population-scale monitoring of shifts in human needs can inform policies and recovery efforts for current and anticipated needs (Section 3.1.7).

Following the examination of population-scale shifts in human needs, the second study [403] characterizes the degree to which different subpopulations shifted their needs in reaction to the pandemic (Section 3.2). In order to capture a holistic view of the changes in digital engagement during the pandemic [347], we structure the analysis according to the five social determinants of health (SDoH) categories defined by the US Department of Health [419], which have been widely used as a holistic framework to describe a wide range of socioeconomic and environmental factors that determine one's health, well-being, and quality of life. We analyze over 57 billion everyday web search interactions during the pandemic across 25,150 US ZIP codes to reveal that the extent to which different communities of internet users enlist digital resources varies based on socioeconomic and environmental factors (Section 3.2.1). For example, we find that ZIP codes with lower income intensified their access to health information to a smaller extent than ZIP codes with higher income. The study also shows that ZIP codes with higher proportions of Black or Hispanic residents intensified their access to unemployment resources to a greater extent while revealing patterns of unemployment site visits unseen by the claims data (Section 3.2.2). Such differences frame important questions on the relationship between differential information search behaviors and the downstream real-world implications on more and less advantaged populations (Section 3.2.3).

The key takeaways and contributions to my thesis from these two studies are summarized in Section 3.3. Both studies were conducted in collaboration with Tim Althoff, Eric Horvitz, and Ryen W. White where I contributed to the conceptualization of the study, the design and refinement of the methodology, the data collection and analysis, and the interpretation of results. The first study was published in WSDM 2021 [402], and the second study was published in Nature Communications [403]. I presented the research conducted in this chapter as a distinguished invited talk at the Web Conference 2021 Web of Health track.

3.1 HUMAN NEEDS FRAMEWORK AND SHIFTING PRIORITIES

Many of the existing studies and datasets of the COVID-19 global pandemic focus on the biomedical and epidemiological aspects of the case and fatality rates, including efforts in detection, infection propagation, therapeutic intervention, and vaccine design, with a gaze fixed on the virus and illness that it causes. Despite the direct focus on mitigating the spread and morbidity of infection [440], pandemic-related policy decisions and investments cannot be made on health information alone. Recent work has called for identifying and understanding the multi-level system of human needs and well-being for pandemic response and recovery strategies [347]. Our goal is to better understand the influences of the pandemic and associated policy decisions on a multitude of human needs, where new insights about shifting needs can guide valuable refinements of policies and motivate the development of new interventions, programs, and investments.

Quantifying human needs across the population is important but challenging, as it requires innovative, ethical, privacy-preserving approaches with fine-grained and broad geo-temporal coverage. A standard way to assess human needs is through survey-based measures [147], which can be costly and time-consuming to conduct at large scales. Surveys are difficult to manage across time and geographies when the desire is to provide fine-grained analyses longitudinally and to be able to understand and react in near real-time. Passively observing human behaviors is another approach, grounded in the fact that physical and psychosocial needs motivate human behaviors to express and fulfill those needs when they are unmet [264]. For example, historical purchase behaviors are used to predict future consumer needs in market research [2], but this approach is also limited to smaller-scale, consumer and commercial interests. E-commerce platforms (e.g., Amazon marketplace) or specialized service providers (e.g., Talkspace, Coursera) may have access to large-scale, real-time analysis of customer behaviors, but they are focused tightly on specific needs. Publicly available social network data (e.g., Twitter), have been used to characterize needs [9, 459], but these studies examine a subset of needs from data that only portrays externalized behaviors. Such fragmentation of data limits the capture of broader expressions and comparisons across a broad spectrum of human needs.

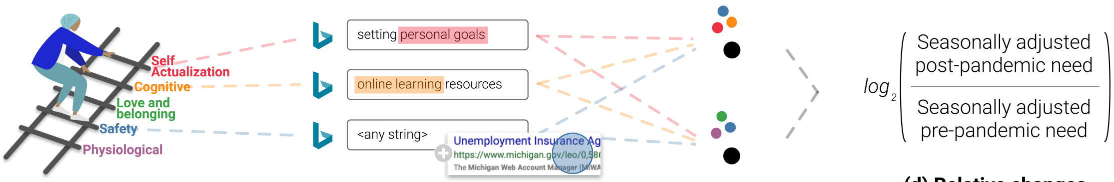
We address these limitations in obtaining signals about human needs by observing the behaviors of people through their everyday interactions with a web search engine. Human behaviors, through which human needs are expressed or fulfilled, often involve seeking information or obtaining tangible support or material items, for which web search has been an integral component. Thus, search logs provide a unique lens into human needs by providing signals about human behaviors in their natural state, at a large scale, and on already routinely collected data.

We propose a computational methodology built on constructs of human needs by Maslow [264, 265] and Max-Neef [269]. We characterize pandemic changes across a broad spectrum of *fundamental human needs*, spanning five broad human needs categories–*Self Actualization*, *Cognitive*, *Love and Belonging*, *Safety*, and *Physiological*–and 79 subcategories. We apply this framework to a dataset of 35+ billion search interactions across 36,000+ ZIP codes in the United States over 14 months (7 months in 2019 and 7 months in 2020) to map search query strings and click interactions to human needs, resulting in over three billion expressions of human needs. We demonstrate how this approach enables the examination of shifts in fundamental human needs based on disruptions induced by the pandemic.

Our contributions include the following:

- We propose a novel computational framework for characterizing a *holistic* view of human behaviors, intents, and unmet needs based on web search logs and human motivational theories (Section 3.1.2).
- We leverage a difference-in-differences approach [234] to quantify the impact of the pandemic and its associated policies on the relative changes in needs while controlling for seasonality and external factors (Section 3.1.2).
- We present the first population-scale analysis across a *holistic* set of human needs during the COVID-19 pandemic in the United States through the use of web search logs (35+ billion search interactions for 14 months on 86% of US ZIP codes; Section 3.1.3-Section 3.1.6).
- We find that search interactions in pursuit of basic human needs (i.e., *Physiological*, *Safety*) have increased exponentially during the pandemic while several higher-level aspirations (i.e., *Self Actualization*, *Cognitive*) have declined (Section 3.1.3, Section 3.1.4).
- We observe geographical differences in how differing statewide shelter-in-place policies are associated with short-term and long-term changes in social and emotional needs (Section 3.1.5).
- We demonstrate that potential barriers to accessing critical resources, in support of people facing unemployment or domestic violence, can be identified through search interactions combined with external data sources (Section 3.1.6).

Our work suggests that signals from web search logs can be used to characterize and monitor over time human needs at a population scale. Our findings also emphasize the importance of tracking broad sets of human needs in combination with other reported measures to identify gaps in our current understanding of challenges and support, measure the impact of policy changes, and design policies and programs that can meet the needs.



(a) Human needs **(b) Expressed needs** **(c) Geo-temporal aggregation** **(d) Relative changes in needs**

Figure 3.1: Illustration of human needs detection framework. (a) Human needs are represented by a ladder according to Maslow's hierarchy of needs to indicate than a person may have multiple needs simultaneously. (b) These needs are expressed through search interactions, which can be categorized through keyword matches and/or subsequent clicks into relevant search result pages. (c) Each search behavior is then aggregated across categories of human needs, time, or geography. (d) To quantify changes in needs, aggregated needs are compared between pre-pandemic and pandemic periods while adjusting for seasonal and query volume variation.

Observation period	14 months (Jan 1-Aug 2 in 2019-2020)
# of queries	35,650,687,581
# of human need queries	3,250,228,644
# of days	428
# of ZIP codes	36,667

Table 3.1: Descriptive statistics for our web search dataset.

3.1.1 Dataset and Validity

Dataset, Privacy, and Ethics

We collected a dataset containing a random sample of deidentified search interactions from the first seven months of the years 2019 and 2020 obtained from Microsoft's Bing search engine. For each search interaction, we collected the search query strings, all subsequent clicks from the search results page (e.g., clicked URL), and the time and ZIP code location of the search interaction. The resulting dataset contains 35+ billion search interactions and represents the web search traffic of over 86% (36,667/42,632) of US ZIP codes associated with at least 100 queries per month so as to preserve anonymity (Table 3.1). All data were deidentified, aggregated to ZIP code levels or higher, and stored in a way to preserve the privacy of the users and in accordance with Bing's Privacy Policy. Our study was approved by the Microsoft Research Institutional Review Board (IRB).

Validation of Data

Considering potential threats to validity, we examined the dataset from three perspectives: coverage of the population, analysis of selection bias, and reliability of the trends.

Analyzing National Representation To understand how much of the US population is represented by the collected data, we obtained demographics data from the Census Reporter API [418]. The demographics of the ZIP codes in our dataset closely matched the US population demographics (Appendix A.1). Although query volumes are not uniformly distributed across these ZIP codes, the vast majority of the ZIP codes are included in our dataset. We leverage location information for our analysis when geographical differentiation is necessary.

Analyzing Selection Bias To understand potential biases in socioeconomic circumstances that would influence the usage of the Bing search engine, we leveraged deidentified client id as a proxy for a unique user to estimate the 'client rate', or how much of the population in a ZIP code is using the Bing search engine. We examined the correlation between the client rate and various demographic factors. While the factors describe some of the variances, none were correlated more strongly than $r{=}{-}0.058$ (% Housing Owned), suggesting that the dataset is not strongly biased towards any single demographic (Appendix A.2).

Reliability of Search Interaction Trends Many Americans use other search engines such as Google. Therefore, we compared search trends for Bing with data available via the Google Trends API for the same time period and for specific keywords in each need category. We found that the search trends are highly similar, with a median Pearson correlation of 0.96 (min=0.45, max=0.98, all $p{<}0.001$). This implies that our findings are not simply an artifact of using one search engine over another (Appendix A.3).

3.1.2 Human Needs Framework

Human Needs Categories

We draw inspiration from Maslow's hierarchy of fundamental human needs [264, 265] to tag each search interaction with one or more of five broad categories of needs. We represent human needs as a ladder to convey that a person may have multiple needs (Figure 3.1a). *Safety* and *Physiological* are considered as 'basic' needs. *Love and Belonging* is often considered to be 'psychological' needs, and *Cognitive* and *Self Actualization* are considered as 'growth' needs, defined in more detail below:

Self Actualization needs are about realizing personal potential, seeking personal growth, and self-fulfillment. Topics include: hobbies; parenting; wedding; talent acquisition; goals; charity.

Cognitive needs are about pursuing knowledge and intelligence through learning and discovering. Topics include: online education, learning materials; educational degrees; cognition; memory; focus.

Love and Belonging are social and emotional needs and include emotionally-based relationships such as friendship, family, dating, sexual intimacy. Topics include: mental health or emotions; social network or activities; relationships, dating, divorce or breakup.

Safety needs stem from our desire to seek order, stability, and protection from elements in the world. Topics include: personal protection; finances; banking; job search; unemployment; housing.

Physiological needs are the basic animal needs such as air, food, drink, shelter, warmth, sex, and other body needs. Topics include: health; food and groceries; basic staples; sleep; transportation.

To understand the nuances of the needs, we further subdivided the five main categories into 79 subcategories. Several researchers independently developed subcategories which were combined and resolved collaboratively through consensus meetings. Appendix A.4 describes this process in detail with the full taxonomy of our need categories, example queries and/or clicked page URLs.

Human Needs Detection

A search interaction can be an observation of the underlying human need in two ways: (1) an expression of a potential satisfier (physical or information) for that need, or (2) a direct expression of the deficiency or satisfaction of that need. For example, a search query for 'bandages' with a subsequent click on 'amazon.com' could indicate a purchase intent that satisfies a *Physiological* need. We require the additional click into one of many e-commerce domains to solidify that this interaction is a purchase intent. Information search about 'online games with friends' could satisfy a *Love and Belonging* need. A need (satisfaction or deficiency thereof) could be directly expressed in experiential statements such as 'I feel depressed' (*Love and Belonging*).

We match each search interaction to a corresponding need subcategory through simple detectors based on regular expressions and basic propositional logic. Each need subcategory could have multiple regular expressions applied to either the query string, the clicked URL, or both, depending on the complexity of the expression and the need subcategory. We arrived at these regular expressions based on our data through several collaborative consensus meetings until we were satisfied with precision and recall (Appendix A.4). Overall, 9.1% of our query samples matched at least one of the need categories. Each search interaction can satisfy multiple human needs [269], so we allowed each search interaction to be tagged by multiple need categories (only 0.32% have multiple tags). We then aggregated matched search interactions across need categories and subcategories, time (e.g., day, week), and geography (e.g., ZIP code, county, state). Figure 3.1 illustrates these steps in detecting and processing of human needs.

Throughout the chapter, we report needs as *expressed* through search interactions and not actual underlying human needs. Some human needs are expressed well by search interactions while other needs are more appropriately expressed through other digital and/or non-digital means. When analyzing *changes* in such search interactions, careful consideration is required to differentiate actual shifts in the underlying human needs (e.g., health-related needs have increased due to COVID-19) from the equally meaningful changes observed in logs based upon shifts from offline to online behaviors (e.g., online grocery purchases due to store closures), and we present our results in both possible contexts.

Framework Validation

Our goal is to detect with high precision a large number of needs across a broad set of categories.

Precision We sampled 1.2 million search interactions that matched at least one need category as a candidate set. From this sample, we randomly chose 100 unique tuples of the search query string and clicked URL (e.g., '15 lb dumbbells' and click on 'walmart.com') for each of the five high-level need categories, for a total of 500 unique tuples representing 1,530 search interactions in our evaluation set. We selected unique tuples to avoid duplication in labeling, but we mapped the labels back to the original 1,530 search interactions to compute precision on the distribution of the source evaluation set. We then collected human labels for each tuple via Amazon's Mechanical Turk, where each tuple could be tagged with none, one, or more of the five needs categories. All labels and predictions have Boolean values with no ranking among needs categories.

Upon inspecting the label quality, we found common systematic label errors such as labeling 'recipe'-related queries as *Cognitive* needs, or 'divorce'-related queries or visits to specific government unemployment sites as *Physiological* needs, where the workers mislabeled the queries according to the definitions we specified in the task detail. Other errors were due to inherent ambiguities in the search. For example, 'rent in florida coronavirus' is tagged as *Physiological* for 'coronavirus' but not as *Safety* for 'rent' because our high precision detector requires more qualified keywords such as 'apartment rent.' Although the worker tagged this as *Safety* (i.e., rent for shelter), the use of the word 'rent' here may not be shelter-related. We took a conservative approach of only correcting definitive label errors and not ambiguous errors, and our evaluation set achieved a precision of 97.2%, using the example-based precision metric defined in [467] for multi-label classification.

Recall Although it is infeasible to ensure a perfect recall across a massive dataset, it is important that we capture a significant number of needs expressions. We find that 9.1% of our search interactions match at least one need category. While this recall is significant and led to more than 3.2 billion detections of needs expressions, we note that high recall is not necessary for an unbiased analysis approach, because we conduct a fair comparison among the outputs of the same detectors across pre-pandemic and pandemic periods. We did not find that the exact expressions of needs varied drastically within our dataset which would indicate any temporal bias. We also investigated whether our need expressions were dominated by a few categories. Clicks to YouTube or Facebook dominated, but still only represented 1% of our dataset. We categorized visits to these sites based on their primary functions (i.e., Facebook for social networking and YouTube for media consumption). We found that our results were robust, whether or not we included these high-traffic sites in our dataset.

Quantifying Changes in Human Needs

Our goal is to quantify the change in human needs during the pandemic relative to the pre-pandemic period. This can be challenging due to the potential confounding effects of yearly seasonal variations, weekly seasonal variations, and variations in query volume over time. Conceptually, we control for yearly seasonal effects through comparisons with the previous year, for weekly seasonal effects by matching the day of the week between both years (i.e., Mon Jan 6, 2020 is aligned to Mon Jan 7, 2019), and by considering relative proportions of the query volume represented by each need over time. Formally, we follow a difference-in-differences methodology [234, 118], commonly used in economics, to account for confounding effects between comparison groups. Finally, our adjusted effect size is the logarithm of the ratio between the two groups. This is effectively the difference-in-differences approach applied to the logarithmic effect sizes and has the advantage of the effect sizes having symmetric properties (i.e., $\Delta(t_1;t_2)$=$-\Delta(t_2;t_1)$) [157, 92]. This step allows for an appropriate comparison of effect sizes across both increases and decreases in need. Our estimate of the *relative change in human need* C between two time periods is defined as

$$C(t_1;t_2,n) = \log_2\left(\frac{E(t_2^{2020},n)}{E(t_1^{2020},n)}\right) - \log_2\left(\frac{E(t_2^{2019},n)}{E(t_1^{2019},n)}\right) \tag{3.1}$$

where $E(t_2^{2020}, n)$ is the expression of the need n at some time t_2^{2020} in 2020 (i.e., after the pandemic declaration) and $E(t_1^{2020}, n)$ is the expression of need at t_1^{2020} (i.e., before the pandemic).

Across all following analyses, we choose the mean daily expression of needs between Jan 6 to Feb 23, 2020 as the 'pre-pandemic baseline', referred to throughout the chapter, and dates on or after Mar 16 as the 'pandemic period' because individual states declared a state of emergency at different times (Feb 29 to Mar 15). We then compute the 95% confidence interval on this multiplicative effect size by using bootstrap resampling with replacement (N=500). We report mean estimates and p-values throughout the text and 95% confidence intervals in all figures and tables where applicable. All time series figures (Figure 3.2, Figure 3.3, Figure 3.5B) indicate the moving average of the daily relative changes, computed from 3 days before to 3 days after.

3.1.3 Temporal Changes in Human Needs

We first consider how human needs change over time across the US in the context of major events surrounding the pandemic. We compute the daily relative change in the expressed needs in comparison to the pre-pandemic baseline, as described in Section 3.1.2, for the duration of our entire dataset, giving us per-day relative changes in all need categories and subcategories. For each inflection point and major national event, we examine need subcategories with the highest relative changes to understand which contributes the most to the overall need.

Elevated Needs and Contributing Subcategories

Figure 3.2 illustrates daily relative changes of needs on a log scale, where zero indicates no change. Overall, we see that all need categories were at elevated rates during March through May relative to the earlier months. A few of the local inflection points correspond to US national events, such as the declaration of national emergency on Mar 13 or the first stimulus checks being deposited on Apr 11.

Physiological needs start to increase first around February (Figure 3.2A), dominated by *health condition related queries* (C=1.46 on Feb 29) and subsequently by *toilet paper purchase* and *health measurement equipment purchase* (C=1.14, 0.76 on Mar 6 respectively; Figure 3.2B). Around

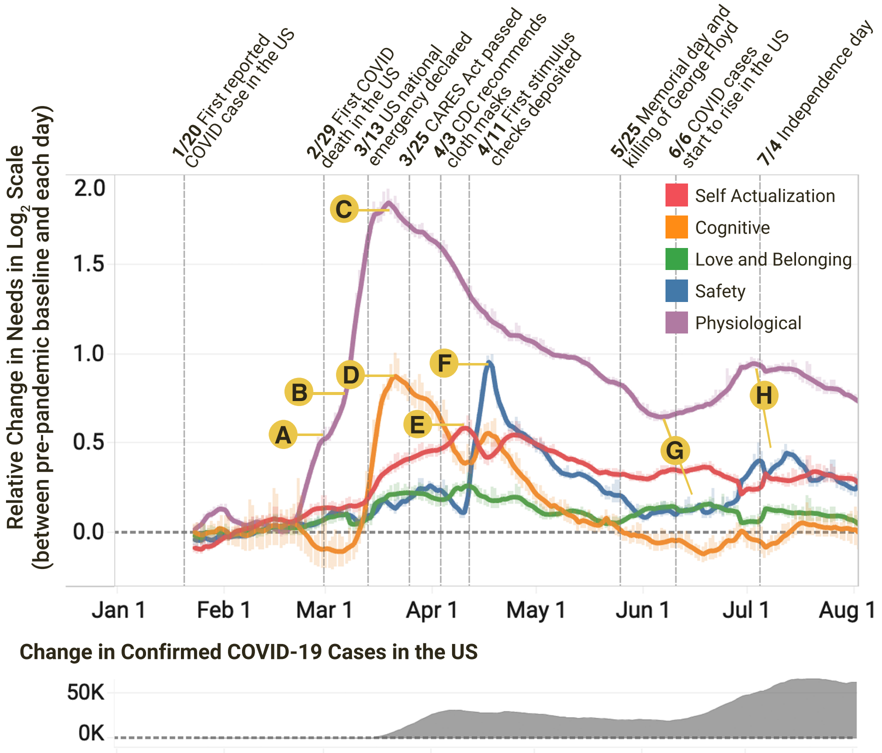


Figure 3.2: Daily relative changes in needs for each human need category throughout the pandemic. Relevant events are annotated vertically at the time of occurrence. Changes in confirmed cases in the US from the 'Bing COVID-19 Tracker' (http://www.bing.com/covid) are displayed below. See Section 3.1.3 for discussion on A-H.

Mar 16, *Physiological* needs peak at over 3.8 times the baseline ($2^{1.91}$; Figure 3.2C). Following the national emergency declaration (Mar 13) and mandated lockdowns (first on Mar 21), we see a sharp increase in *Cognitive* needs (Figure 3.2D), dominated by *educational site visits* and *online education queries* (C=1.97, 1.42 on Mar 23). *Self Actualization* needs peak around Apr 11 (Figure 3.2E), dominated by *cooking site visits* and *cooking related queries* (C=1.73, 1.15), and *online social activities queries* and *social technology uses* dominate *Love and Belonging* needs (C=1.77, 1.72 on Apr 11). A sharp spike of *Safety* needs can be seen shortly after the first stimulus checks were deposited: *stimulus related queries*, *state unemployment site visits*, *COVID-19 protection purchase* dominate *Safety* needs (C=8.17, 5.49, 5.12 on Apr 18; Figure 3.2F). While other needs start to trend downwards or stabilize throughout much of May-July, *Physiological* needs increase for a second time with additional interests in health conditions, followed by *Safety* needs with queries related to economic stimulus and loans (Figure 3.2G), which aligns with the rise of COVID-19 cases in the US around Jun 6.

Shifting of Needs

Based on the severe health impacts of the COVID-19 pandemic, we expected to see and confirmed that *Physiological* needs dominated throughout our dataset, as COVID-19 is still a major US public health issue at the time of writing. At a glance, we see two instances of the surge in *Physiological* needs, followed by a subsequent increase in *Safety* needs. As *Physiological* concerns rise, public health responses (e.g., business closures or restrictions) could potentially induce instabilities in *Safety* needs, and this observation needs to be further investigated. We see basic needs expressed before other needs consistent with the hypothesis by Maslow [264] and observations by others [407]. We also expected to see a decrease in the expression of growth needs (*Self Actualization* and *Cognitive*) as people's attention shifts toward basic needs. However, both *Cognitive* and *Self Actualization* needs increased overall, with the increase in *Cognitive* needs being more temporary and *Self Actualization* being more sustained. Despite health and economic concerns, interests in recreational activities or hobbies (e.g., cooking, gaming) contribute to this steady 23% ($=2^{0.3}-1$) increase in *Self Actualization* needs. Further research into the temporary nature of *Cognitive* needs and the long-term impact of such sustained interest in *Self Actualization* is necessary.

We see that the peak in *Physiological* needs occurs around four weeks before the peak in *Safety* needs (Figure 3.2C, F), while the second set of peaks is a few days apart (Figure 3.2H). This could be an indication of phenomena like resilience or endurance from economics and disaster management that requires further examination [263, 142].

3.1.4 Significant Changes in Human Needs

Next, we examine individual need subcategories that present the largest increase or decrease in search expressions, possibly due to the pandemic's impact. To explore these two ends of

Need	**Need Subcategory**	$C_{\textsf{mean}}$	$C_{\textsf{max}}$	$2^{C_{\textsf{max}}}$-1
Phys	Toilet paper purchase	6.11±0.08	7.00	12691.1%
Safe	Stimulus related queries	5.69±0.03	8.17	28601.9%
Safe	Unemploy. related queries	4.88±0.02	5.72	5156.4%
Safe	State unemploy. site visits	4.81±0.04	6.26	7585.1%
Safe	COVID-19 prot. purchase	4.67±0.03	5.57	4634.0%
Phys	Health meas. equip. purchase	3.43±0.04	4.56	2257.2%
Phys	Health cond. related queries	3.12±0.01	3.54	1065.6%
Phys	Food assist. related queries	2.62±0.03	3.12	771.8%
Phys	Grocery related queries	2.05±0.01	2.54	480.8%
Phys	Food delivery queries	1.84±0.02	2.26	379.7%
L&B	Online social activity queries	1.77±0.09	2.98	688.4%
Phys	Food delivery site visits	1.7±0.01	2.26	379.3%

Table 3.2: Top 12 need subcategories with the largest *increase* in the mean relative change in need within the initial 4 weeks of the pandemic with 95% confidence intervals, the maximum relative change in the dataset, and maximum percent change.

Need	**Need Subcategory**	$C_{\textsf{mean}}$	$C_{\textsf{min}}$	$2^{C_{\textsf{min}}}$-1
SA	Wedding related purchase	-1.49±0.03	-1.76	-70.4%
SA	Wedding site visits	-1.25±0.02	-1.56	-66.2%
Cog	Edu. degree related queries	-0.87±0.06	-1.09	-53.1%
Safe	Housing related queries	-0.71±0.05	-1.09	-53.2%
Safe	Job search related queries	-0.65±0.03	-0.91	-46.7%
Safe	Job search site visits	-0.61±0.02	-0.95	-48.1%
Phys	Apparel purchase	-0.60±0.01	-0.84	-44.1%
SA	Outdoor related queries	-0.59±0.01	-1.07	-52.3%
SA	Life goal related queries	-0.57±0.09	-1.23	-57.2%
Safe	Domestic violence queries	-0.54±0.04	-0.97	-49.0%
Safe	Rental related queries	-0.53±0.06	-0.82	-43.5%
L&B	Divorce related queries	-0.49±0.03	-0.93	-47.3%

Table 3.3: Top 12 need subcategories with the largest *decrease* in the mean relative change in need within the initial 4 weeks of the pandemic with 95% confidence intervals, the minimum relative change in the dataset, and minimum percent change.

the spectrum, we compute the mean relative changes in needs during the initial four weeks of the pandemic period (Mar 16 to Apr 12) compared to the pre-pandemic baseline, as described in Section 3.1.2. We then examined the top 12 need subcategories with the largest increase or decrease in the relative change.

Heightened Physiological *and* Safety *Needs*

Table 3.2 shows that 11 need subcategories with the most increase fall under *Physiological* and *Safety* needs as seen from the temporal trends in Section 3.1.3, and one (*online social activities queries*) belongs to the *Love and Belonging* need category. *Toilet paper purchase* reached a maximum increase of 127 times the pre-pandemic baseline (Mar 16). Recall that these are not just queries containing 'toilet paper', but purchase intents as indicated by subsequent clicks to e-commerce sites (Section 3.1.2). Such a high level of interest in toilet paper is commonly attributed to panic buying due to the supply scarcity [167] and media coverage [148]. *Stimulus related queries*, including general terms like 'loan forgiveness,' reached an even higher maximum increase of 286 times the baseline (Apr 18) and is sustained at that high level through July, reflecting the magnitude of the pandemic's impact on the US economy.

When we examine the daily trends, we see that *COVID-19 protection purchase* exhibits a small peak (C=1.0 on Jan 29) after the first reported COVID-19 case (Jan 20), and the needs quickly escalate from Feb 20 (Figure 3.3B), at least three weeks earlier than other needs (Figure 3.3A,C,D) that do not escalate until the national emergency declaration. A subsequent peak on Apr 3 (C=5.6) coincides with CDC's updated recommendation on cloth-based mask use. We also find that indicators of social-economic instabilities such as *unemployment site visits* and *food assistance related queries* still have not returned to their baseline levels (Figure 3.3A,C), arguably because the pandemic is still in effect. *Online social activities queries* follow a similar pattern (Figure 3.3D), reflecting the need to satisfy lock-down-induced social isolation through online services, but it also raises the question of potential permanent shifts in ways of satisfying social needs. Our findings revealed that only a few of these needs have returned to the pre-pandemic baselines while many of them are sustained at elevated rates.

Shifts Away from Positive Outlooks

Table 3.3 shows the most decrease in the expression of several *Self Actualization* and one *Cognitive* need subcategories. Specifically, indications of *Self Actualization* needs for partnership have

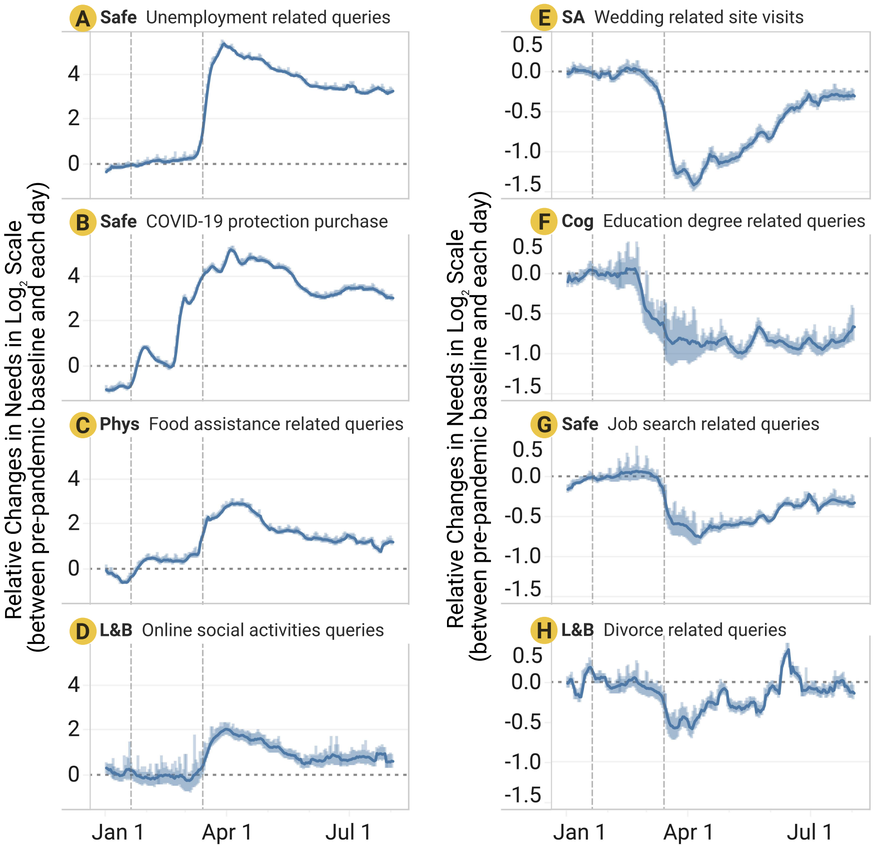


Figure 3.3: Daily relative changes in select needs with the largest increase (A-D) and the largest decrease (E-H) in need. Vertical bars denote the first reported US COVID case (Jan 20) and the US national emergency declaration (Mar 13).

declined by more than 64% (=$2^{-1.49}-1$) of their baseline throughout the typical US wedding season around Spring and early Summer, which is expected given restrictions on large gatherings. In addition, needs that are typically associated with growth, positive outlook, or new opportunities have taken a large toll. In *Self Actualization*, queries about life goals also see a large decline. In other need categories, needs expressed by *educational degree related queries*, *job search related queries*, *job search site visits*, or *housing related queries* have declined by over 34% (=$2^{-0.61}-1$) of their baseline.

Upon inspection of daily trends, expressions of forward-looking needs have decreased and remained below the pre-pandemic baseline (Figure 3.3E,F,G). The sustained decline in *job search* *related queries* (Figure 3.3G) juxtaposed with near 30 times increase in unemployment needs (Figure 3.3A) is troubling evidence of the declining labor force as seen in other studies [91]. Indications of growth interests in educational degrees or life goals have not recovered (Figure 3.3F). These results combined with heightened *Physiological* needs suggest a shift of focus away from individual growth. *Divorce related queries* exhibited a maximum of 47% (=$2^{-0.93}$; Figure 3.3H) decline, possibly reflecting the challenges that families face in proceeding with divorce during the lock downs [233]. Therefore, the underlying mechanisms for these shifts, the long-term impact of the lack of growth needs, and support for relationships should be further studied. See Appendix A.6 for corresponding figures on all 24 subcategories.

3.1.5 Geographical Differences in Needs

We shift our attention to how the pandemic and its related policies *differentially* influence local *subpopulations*. We use a set of statewide policies[1] readily available through the COVID-19 US State Policy Database [323]. We examine the impact of the shelter-in-place mandate (its duration and effective date) on social-emotional and relationship needs (i.e., eight *Love and Belonging* and two wedding-related *Self Actualization* subcategories). We hypothesized that social isolation induced by longer shelter-in-place mandates generates more expressions of social-emotional needs.

We used the date on which the shelter-in-place mandate was enacted and relaxed or lifted for 38 states to derive the mandate duration[2]. We compute two relative changes in needs for each state. First, to understand the *short-term* impact of the shelter-in-place mandate, we compute the relative change in needs for *two weeks after* each state's shelter-in-place mandate compared to the *one week before* the mandates[3]. Second, to understand the *long-term* impact of the mandate, we compute the relative change in needs between the *pre-pandemic baseline* and the *last four weeks* of our dataset (Jul 6 to Aug 2), which is at least two weeks after the last state lifted its mandate (Jun

---

[1] Although our dataset allows ZIP code level analysis, a comprehensive list of local policies across the US is difficult to obtain at ZIP code, city, or even county levels.

[2] Only 38 states had the start and end dates in the dataset as of August 2020.

[3] We chose one week before the mandate because it is the maximum number of full weeks after the declaration of national emergency and before the earliest start date (Mar 21) for any shelter-in-place mandate.

19). To quantify the potential impact of these mandates, we ran a Pearson correlation analysis (1) between the start date of the shelter-in-place mandates (i.e., ISO day number) and the short-term relative changes in needs, and (2) between the duration of the mandates and the long-term relative changes in needs.

Early Adjustments to Social Needs

The start date of shelter-in-place mandates ranged between Mar 21 and Apr 7. We find that the relative changes in *online social activities queries* (r=-0.53, p<0.001), *wedding site visits* ((r=0.48, p=0.002), *wedding related purchase* (r=0.43, p=0.006), and *mental health resource site visits* (r=0.47, p=0.003) needs were most correlated with the start date. People from states that have earlier shelter-in-place mandates expressed significantly reduced interest in weddings and mental health site visits (Figure 3.4A) and significantly more need for online social activities (Figure 3.4B). For example, in the two weeks after the mandate, people in New Jersey (mandate on Mar 21) sought online mental health resources 25.7% less than the week before the mandate, while people in South Carolina (mandate on Apr 7) sought those resources 13.4% more than the week before.

Long-term Social Impact

When we examine the long-term changes in expressed needs, we find that shifts in *negative mental health experiential queries* (r=0.42, p=0.010, Figure 3.4C), *wedding site visits* (r=-0.37, p=0.022, Figure 3.4D), and *wedding related purchase* (r=-0.35, p=0.033) were most correlated with the duration of sheltering and closures. The duration of the shelter-in-place mandates was highly correlated with the start date (r=-0.62, p<0.001), indicating that people faced with earlier mandates are also impacted by longer mandates. As we discovered above, these people likely suppressed their needs for weddings or mental health support early during the pandemic. At the same time, as wedding needs slowly recover to the pre-pandemic baseline (Figure 3.3E), those impacted by longer mandates are even slower in their recovery of wedding needs and are more likely to express negative mental health issues. For example, people in Mississippi (24 days of shelter-in-place) expressed 33.2% *less* negative mental health experiences than before the pandemic

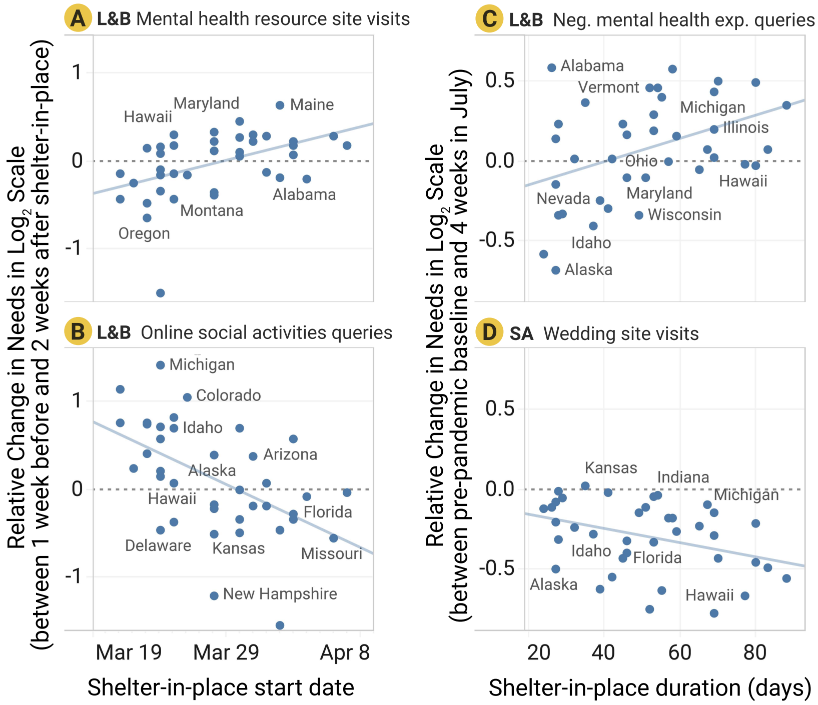


Figure 3.4: Relationships between shelter-in-place start date and *short-term* relative changes in mental health and online social activities needs between 1 week before and 2 weeks after shelter-in-place mandates (A-B), and between shelter-in-place duration and *long-term* relative changes in mental health and wedding needs between pre-pandemic baseline and 4 weeks in July (C-D).

while people in Oregon (88 days) expressed 27.2% *more* negative mental health experiences than before (Figure 3.4C). Others attribute this increase in negative emotions during the pandemic to the shift to basic needs [79]. Per-state analysis of this shift along with the differential prevalence of COVID-19 may be necessary to understand the mechanisms for this increase in negative emotions and to provide appropriate social-emotional support.

3.1.6 Gaps between Expressed and Reported Needs

As we have demonstrated so far, there are many needs that are well expressed by web search interactions such as purchasing goods online, looking for health information online, or accessing economic assistance through government websites. Our final analysis examines a gap between how web search facilitates the expression of these needs and the reported fulfillment of these needs

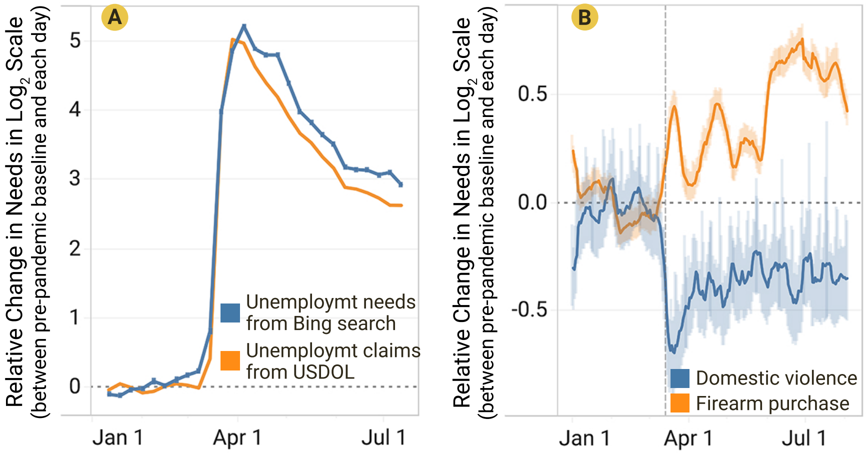


Figure 3.5: Weekly relative changes in unemployment needs from Bing search and in unemployment claims from the US Department of Labor compared to pre-pandemic baseline (A), and daily relative changes in domestic violence and firearm purchase needs compared to pre-pandemic baseline (B).

to discover potential barriers to accessing critical resources. To demonstrate how our approach allows for exploring these barriers, we focus on two need subcategories: unemployment and domestic violence.

Expressed Unemployment Needs vs. Reported Claims

We obtained weekly, seasonally adjusted initial unemployment claims for 2020 from the US Department of Labor[4] and computed the relative changes in unemployment claims using the same approach as in Section 3.1.2. We compare the change in reported unemployment claims with the change in expression of unemployment needs in search and find that these two changes are closely aligned (Pearson r=0.996, p<0.001) (Figure 3.5A). Despite policies that extend unemployment eligibility for up to 26 weeks during the pandemic, we see that the expression of unemployment needs remains at 25% higher than the reported claims since April. This discrepancy corroborates with known issues with the unemployment benefits: many people were confused about the benefits (e.g., job search requirement) and were being denied, having to search for more information or

---

[4] https://oui.doleta.gov/unemploy/claims.asp

file multiple applications [367, 46]. Although our analysis excludes those who use traditional methods for filing unemployment claims (e.g., mail, phone), our results indicate that web search is a critical resource that facilitates unemployment needs and highlights a potential gap in satisfying unemployment needs which requires further investigation.

Expressed Domestic Violence Needs vs. Reports

One of the dire consequences of the COVID-19 pandemic is an increased risk of domestic violence due to shelter-in-place mandates exacerbated by physical, financial, and social-emotional stressors and increased alcohol consumption at home [75, 239]. Our analysis shows that the expression of needs for *domestic violence queries* has dropped by nearly 36.7% (=$2^{-0.66}-1$) since the pandemic and stabilized at around -15.9% (=$2^{-0.25}-1$, Figure 3.5B) below pre-pandemic levels. Interests in *firearm purchase* have increased by over 40% (=$2^{0.5}-1$), corroborating other reports of increased gun sales and gun violence and worrisome links to fatal domestic violence incidents [404, 177, 75].

A similar decrease in domestic violence-related metrics in March is seen by national hotlines[5]. Some reports point to potential underreporting by victims unable to reach out for help under constant surveillance by the offenders at home and fearful of exposure to the virus at a shelter [68, 383]. Although the expansion of online resources is critical to addressing the rise in domestic violence [255], many studies highlight the need to recognize control tactics that prevent access to these digital resources [344]. The sustained decrease in domestic violence queries seen in our data may indicate such barriers or shifts in access methods for these online resources. The complexities of domestic violence are evident in the inconsistencies across many data sources, resource media, and contexts. Our results frame an urgent question that needs to be resolved through expertise in domestic violence and social work organizations and through a careful combination of multiple data sources to identify the underlying explanations for our findings.

---

[5] https://crisistrends.org/; https://www.thehotline.org/

3.1.7 Discussion

We presented a computational methodology based on theories of human needs to quantify the effects of the pandemic and its related events on web search interactions. Although all need categories had elevated changes in needs, we found that basic need subcategories were elevated the most while growth-based need subcategories, indicative of positive outlooks in life, were subdued (Section 3.1.3, Section 3.1.4). We also found that earlier and longer shelter-in-place mandates may come with an unintended impact on mental health needs (Section 3.1.5). We used unemployment and domestic violence-related queries to demonstrate how our methodology could help expose gaps between the expression and fulfillment of needs–and frame questions and directions for urgent investigation (Section 3.1.6).

Limitations

We cannot use our method to make causal claims. If other major concurrent events (e.g., Black Lives Matter) have significantly influenced human needs during the observations window, we are not able to distinguish between needs caused by the pandemic or these events. However, our analysis clearly presents changes in the expression of human needs which can be important on its own for understanding rising needs and considering factors for the design of specific interventions. Our analysis relies on heuristics and correct inference of needs from search interactions. But our framework is theory-driven, comprehensive, and achieves 97% precision across five main need categories. Further fine-tuning of the framework should be motivated by precision-recall requirements of individual analysis goals. Although *expressions* of needs may not reflect the actual underlying human needs, the methodology serves as a useful detector for needs, as we have demonstrated through its corroboration with several external findings, and combining our methodology with additional data sources and interdisciplinary collaboration can further the understanding of changes in human needs during global crises. In contrast to traditional methods of measuring needs, our approach can operate at population scales relying solely on data that is already routinely collected by web search engines, and thus allows for effective, retrospective, and fine-grained studies with observation periods of over one year.

Implications: Resilience and Vulnerability

Our work has demonstrated major shifts in a spectrum of human needs during the pandemic, echoing the need to understand the societal, economic, and psychosocial effects of these events and the underlying system of human needs. Such understanding is essential for guiding programs around outreach and support of people during the current pandemic and preparing for future pandemic responses and recovery efforts [347]. We observed that some needs have returned to their baseline levels while others are sustained at elevated (e.g., health and economic instability) or suppressed (e.g., job search, educational degree) levels (Figure 3.3). As we saw in Section 3.1.3, temporal variations on how and when certain needs arise in response to major disruptions could indicate a degree of psychological and economic resilience [263, 142]. Understanding how well a region can endure disruptions is crucial for balancing health risks against societal risks, especially for vulnerable populations who are severely impacted by the pandemic [88]. Our framework could be used to quantify how much a community has been and might be able to endure social and economic distress. For example, a region may be able to withstand prolonged business closures before loan forgiveness needs surge, while another region may show an immediate need for financial stimulus. In addition, when coupled with demographic variables, our framework could be used to examine disparities in needs which may help with understandings of the disparate impacts of blanket policies on the most vulnerable populations.

Implications: Preparedness and Resources

Our work revealed strong correlations indicative of a long-term mental health impact of longer shelter-in-place mandates, suggesting that states with longer duration of shelter-in-place mandates may need to provide additional support for social-emotional well-being (Figure 3.4). We also saw potential barriers to accessing critical resources for unemployment and domestic violence; a decrease in domestic violence queries in spite of conflicting anecdotes should raise questions and an alarm (Figure 3.5). These examples suggest that expressions of needs through web interactions could highlight resource deficiencies or barriers. As regions prepare recovery efforts from COVID-19 or make plans for future disruptions, our methodology can be harnessed to monitor or anticipate a spectrum of needs at various geo-temporal granularities; this could help reveal impacts of the pandemic (e.g., mental health [103]) and cater the appropriate interventions that meet individual and societal needs.

Future Directions

We see this work as a step towards achieving a more holistic understanding of the multiple influences of the pandemic and associated policies and programs. While we focused on retrospective analyses, our approach has the potential to be used in near real-time to monitor human needs and support future policy decisions. Understanding the disparate impacts of the pandemic and its policies on a full spectrum of human needs, especially for vulnerable populations, is critical for designing response and recovery efforts for major disruptions. Our work highlighted future research opportunities and calls for action and collaboration across epidemiology, economy, social sciences, risk management, and more. We look forward to further refinement of the methods presented and hope the work will encourage discussions on observing and addressing changes across a broad spectrum of human needs during a global crisis.

3.2 DISPARATE IMPACT AND BURDENS OF THE PANDEMIC

Socioeconomic and environmental factors play a significant role in the health and well-being of individuals and communities [140, 455, 450]. Despite pandemic-driven efforts to close the long-term and emergent health equity gap [455], studies during the COVID-19 pandemic have demonstrated that socioeconomically and environmentally disadvantaged subpopulations have been disproportionately and negatively affected by the disease [88, 463, 81], with threefold higher infection rates and sixfold higher death rates in predominantly Black US counties than in white counties [458]. In recent decades, digital access has also gained attention as an important factor modulating health outcomes, as individuals harness the internet to seek health information and to access healthcare services (i.e., telehealth, online pharmacy) [165]. During the COVID-19 pandemic, digital engagement in resources across health, educational, economic, and social needs grew in importance because of lockdown mandates, social isolation, and economic burdens [402, 39, 20] as well as due to internet-based communication methods employed by public institutions, such as the online dissemination of COVID-related information by the World Health Organization [39].

Unfortunately, disparities in digital access also reflect socioeconomic and environmental dimensions of variation [143]. The most basic form of digital inequality, the so-called first-level digital divide, manifests itself as the difference between adequate and inadequate digital infrastructure and devices (i.e., access to technology or the quality of access) [423]. Digital inequalities also manifest themselves as the differences in the usage of digital technologies and skills relevant to the usage of digital technologies, the so-called second-level digital divide [172, 174].

In this study, we harness the centrality of web search engines for online information access to observe the second-level digital divide at population scales. We conduct a retrospective and longitudinal observational study using 55 billion everyday web search interactions across multiple devices and 25,150 US ZIP codes during the COVID-19 pandemic. In our work, instead of focusing narrowly on a single topic, we aim to examine a spectrum of broader information domains to capture a holistic view of the changes in digital engagement during the pandemic [347]. Therefore, we structure our analysis according to the five social determinants of health (SDoH) categories defined by the US Department of Health [419], which have been widely used as a holistic framework to describe a wide range of socioeconomic and environmental factors that determine one's health, well-being, and quality of life. We apply the SDoH framework to search data to quantify the changes during the pandemic in how offline exclusion (e.g., lack of sufficient economic resources, lack of health insurance) relates to changes to existing digital exclusion (e.g., reduced participation in online banking or eHealth).

The differential digital engagement patterns we present in this study have real-world downstream implications. Most recently, the third-level digital divide has been conceptualized as the differential ability to translate the use of digital technologies into favorable outcomes, particularly leading to negative downstream outcomes in offline realms such as occupational pursuits, healthcare, and social networking [422, 74]. For example, the digital footprint gap in the usage of information and communication technologies (ICTs) has been shown to surface during childhood and the entire life course along offline axes of socioeconomic status (SES). As a result, they may wind up with smaller social networks and limited employment opportunities [338]. Furthermore, even after controlling for internet access, those from higher SES or higher digital literacy integrate digital resources into their lives and use the internet for more "capital-enhancing" activities that are likely to result in more upwards mobility in the offline world [174, 173, 74, 322]. Just as the social, economic, cultural, and personal offline resources can affect engagement in the corresponding digital fields, digital exclusion and the lack of engagement in digital resources can lead to negative offline consequences [180] across the range of downstream outcomes in the domains of health [356, 165, 338], education [69], and employment [226, 115]. Therefore, it is important to observe digital behaviors across subpopulations and scrutinize the role of digital inequalities in our society. In addition, disadvantaged subpopulations are already at a higher risk of COVID-19 infection and mortality with heavier pandemic-induced socioeconomic burdens, such that it is critical to ensure that digital inequalities do not exacerbate the disparate impacts of the pandemic even further [39].

In this study, we contribute empirical evidence for the second-level digital divide by quantifying how the changes during the pandemic in how offline exclusion (e.g., lack of sufficient economic resources, lack of health insurance) relates to changes to existing digital exclusion (e.g., reduced participation in online banking or eHealth). Unlike prior studies that are limited to interviews, surveys, or specialization in domains or locations, we provide near real-time, population-scale analysis across many different information domains to reveal naturalistic digital engagement patterns uniquely seen through search data. Our analysis provides unique findings that are unobserved by other data sources. For example, we observed a surge in unemployment site visits in August 2020 that are not captured by the unemployment claims data. We also observed differential uses of pandemic-relevant online resources that span across health information, learning, and food delivery. Most importantly, we demonstrate a disproportionate change in a community's use of these digital resources across several socioeconomic and environmental offline factors. These differences are significant when put into the context of the bidirectional nature of digital and offline exclusion. Although the SDoH factors and outcomes reviewed in our analysis are generally not modifiable (e.g., race) or difficult to modify (e.g., income), our findings highlight specific at-risk populations for whom to target shielding or interventions (e.g., stimulus checks to target low-income populations).

3.2.1 Methods

In this section, we describe our data set and analysis methods. At a high-level, our dataset includes anonymized search queries to the Bing search engine and subsequently clicked web site URLs from those queries. Each search interaction is classified into the categories of health, education, economic assistance, and food access that cover a broad range of critical resource needs (Table B.4). We link the search interactions from each United States ZIP code to their respective per-ZIP code census variables that broadly cover five SDoH categories: (1) Healthcare Access and Quality (through health insurance coverage), (2) Education Access and Quality (through educational attainment level), (3) Social and Community Context (through proportions of the population represented by different race/ethnicity), (4) Economic Stability (through income and unemployment rate), and (5) Neighborhood and Built Environment (through population density and internet access). We divide our dataset according to these SDoH factors and compare the magnitude of change in search behaviors between two ZIP code groups during the pandemic, where a larger observed difference in the magnitude of change in search behaviors could indicate that one group's response to the pandemic is more significant than the other in the level of interest in online information (e.g., health, unemployment) or in accessing online resources (e.g., online remote learning).

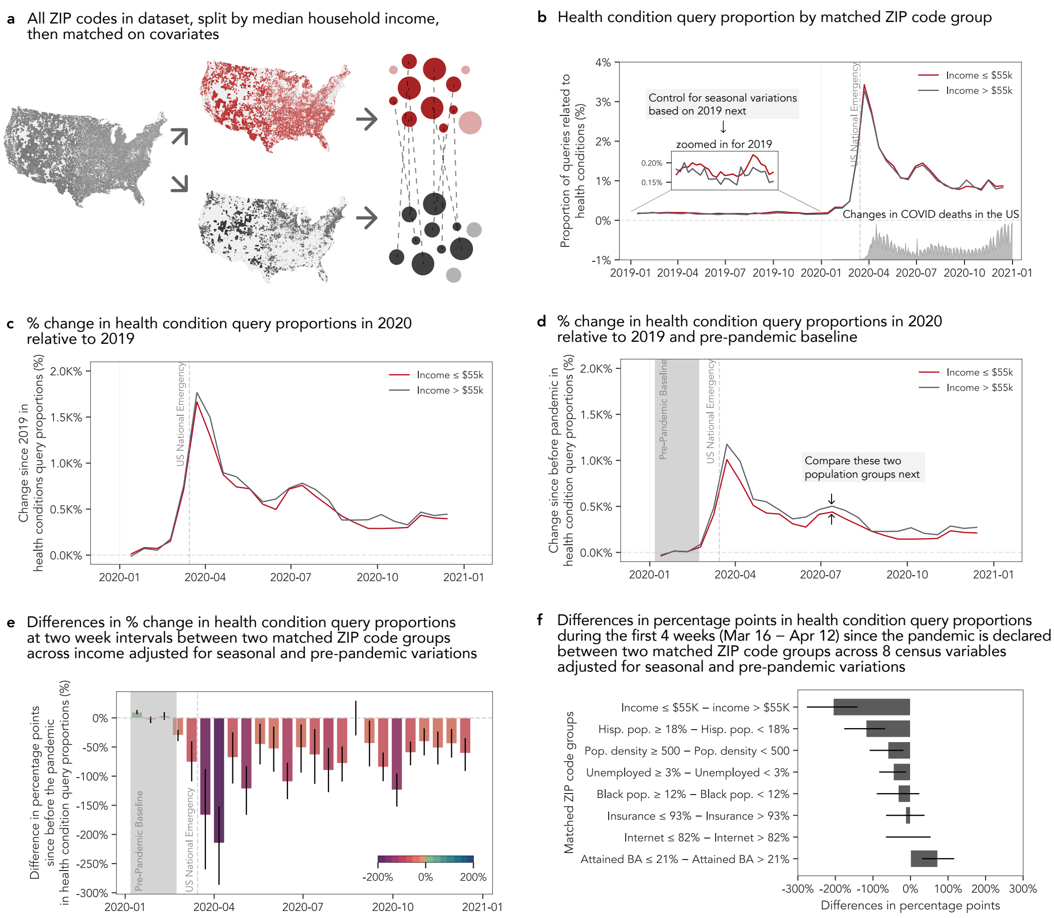
Figure 3.6 illustrates the analysis process for online health information access. For example, we split our ZIP codes into low and high income groups (below and above $55,000 median household income) and compare the magnitude of change in health condition information queries (Fig. 3.6a). To disentangle the confounding effects of SES and race/ethnicity proportions on behaviors and health [5], we compare changes in search behaviors on matched pairs of ZIP codes that are highly similar across these potentially confounding factors (Methods). We isolate the relative changes in search behaviors that occur concurrently with the pandemic using difference-in-differences approach [234], adjusting for yearly and weekly seasonality and for pre-existing, pre-pandemic disparities in query volume (Fig. 3.6b-d, Methods). Thus, we measure the disparate intensification or attenuation of search behaviors during the pandemic between the two ZIP code groups delineated by their distribution in a single SDoH factor (Fig. 3.6e). Finally, we apply the same process across all SDoH factors (Fig. 3.6f, Methods).

The rest of this section goes into the data and the analysis process in detail.

Data Set and Study Population

Our source dataset consists of a random sample of 57 billion de-identified search interactions in the United States from the years 2019 and 2020 from Microsoft's Bing search engine. Each search interaction includes the search query string, URLs of all subsequent clicks from the search result page, timestamp, and ZIP code. We excluded search interactions from ZIP codes with less than 100 queries per month so as to preserve anonymity. Our search dataset intentionally includes both desktop and mobile Bing search interactions in order to capture both search query sources. Although the quality of access, especially through different device types or device specifications, has been highlighted as another important factor in recent digital divide research [423], analysis of the differential search behaviors across device types is outside the study's scope. All data were deidentified, aggregated to ZIP code levels or higher, and stored in a way to preserve the privacy of the users and in accordance to Bing's Privacy Policy. Our study was approved by the Microsoft Research Institutional Review Board (IRB).

While many Americans use other search engines such as Google, Bing's query-based market share is estimated to be about 26.7% according to Comscore data [96]. We focused on query-based metrics for estimating search market share because it captures end-users' interaction with the search engine, including queries that may not have resulted in site visits. Click share, on the other hand, captures only search-driven traffic to a subset of websites that are instrumented with custom code. To understand the validity of relying solely on Bing search data, we compared Bing and Google queries for matched categories longitudinally and found that the search trends are highly



a All ZIP codes in dataset, split by median household income, then matched on covariates

b Health condition query proportion by matched ZIP code group

c % change in health condition query proportions in 2020 relative to 2019

d % change in health condition query proportions in 2020 relative to 2019 and pre-pandemic baseline

e Differences in % change in health condition query proportions at two week intervals between two matched ZIP code groups across income adjusted for seasonal and pre-pandemic variations

f Differences in percentage points in health condition query proportions during the first 4 weeks (Mar 16 – Apr 12) since the pandemic is declared between two matched ZIP code groups across 8 census variables adjusted for seasonal and pre-pandemic variations

Figure 3.6: Quantifying disparities in online health information access. (a) 25,150 ZIP codes above and below \$55,224 median household income are matched to control for other confounding covariates (see Methods). (b) The proportion of queries relating to a collection of health conditions increases dramatically to over 3% around the time the US national emergency was declared. (c) Seasonal and weekly variations are accounted relative to 2019. (d) After accounting for pre-pandemic baseline (relative to January 6 - February 23, 2020, shaded in gray), we isolate the percent change in health condition query proportions introduced during the pandemic where the differences between high- and low income groups start to emerge. (e) We observe that low income ZIP codes experienced almost 200% less change in health condition queries compared to that of the high income groups right after the US national emergency is declared. (f) The same matching-based comparisons are performed across all SDoH factors during the first four weeks since the declaration of the pandemic in the US. In (e) and (f), data are presented as mean values, and error bars indicate 95% confidence intervals via bootstrapping (n=500).

correlated (Pearson $r = 0.86$ to $r = 0.98$, Figure B.3). Our search ZIP code data is provided by a proprietary location inference engine, with added accuracy improvements to standard reverse

IP lookup databases from contextual and historical information, but such estimation is still an approximation. Our study also assumes that the demographics of the search users in a ZIP code reflect the demographics of the population within a ZIP code. However, search users generally trend towards a more white, richer, and older population. It is difficult to accurately characterize the population base without third-party services such as Comscore data [96], which may have its own limitations and biases. Our analysis of a proportion of user demographics using such data confirms that Bing data tracks the US population reasonably well.

ZIP Code Level Data One of our goals is to characterize the role of socioeconomic and environmental factors on digital engagement outcomes. Unfortunately, data that combines individual-level search interactions with each individual's socioeconomic and environmental characteristics at the US national scale does not exist, is difficult to capture, and invites privacy concerns. Instead, we use ZIP codes as our geographic unit of analysis. ZIP code level analysis can be limited because it cannot describe each individual living in those ZIP codes. However, ZIP code level analysis can scale to nontrivial population sizes and has been repeatedly recognized and leveraged in population-scale and local/neighborhood-level research [14, 16, 204, 156, 10, 371, 77]. ZIP code level analysis also enables accounting for well-known issues associated with residential segregation and socioeconomic disparities [143, 449]. We leveraged the available ZIP code level American Community Survey estimates using the Census Reporter API [418] in order to characterize the ZIP codes in our dataset.

Census Variables and Search Categories We chose a set of census variables to delineate ZIP code groups as well as search categories to examine digital behaviors. Figure B.1 illustrates our full choice of census variables and search categories.

The social determinants of health (SDoH) have been widely used as a holistic framework to describe a broad range of socioeconomic and environmental factors that determine one's health, well-being, and quality of life. In recent years, the SDoH has also been referenced in relation to the digital divide; digital literacy and internet access are referred to as "super determinants of health" as they relate to all social determinants of health [373]. Just as Helsper [180] theorized the corresponding digital and offline fields, looking at variables from both offline and digital aspects of the social determinants of health is critical in understanding digital disparities. Because of the multidimensional nature of socioeconomic status and its association with health and well-being outcomes, it is important to include relevant socioeconomic factors [58]. Therefore, our choice of census variables and search categories are largely influenced by the SDoH framework defined by the US Department of Health [419].

We considered multiple socioeconomic factors including race, income, unemployment, insurance coverage, internet access, educational attainment level, population density, age, gender, Gini index, homeownership status, citizenship status, public transportation access, food stamp, and public assistance. We did not include some of the factors when they were highly similar to already included factors (e.g., % below the poverty level is correlated to median household income, Pearson $r = -0.624$). In the end, we included eight census variables that represent all five categories of SDoH to cover a broad range of socioeconomic and environmental factors.

Under Healthcare Access and Quality, we included the percentage of the population with health insurance coverage (Table B27001). Under Education Access and Quality, we included the percentage of the population that attained a Bachelor's degree or higher (Table B15002). Under Social and Community Context, we included the percentage of the population of Hispanic origin (Table B03003) and the percentage of the population with Black or African American alone (Table B02001). Under Economic Stability, we included the median household income (Table B19013) and the percentage of the civilian labor force that is unemployed (Table B23025). Under Neighborhood and Built Environment, we included the percentage of the population with a broadband or dial-up internet subscription (Table B28003) and the population density. We computed per ZIP code population density by joining area measurements from ZIP Code Tabulation Areas Gazetteer Files [417] and total population (Table B01003). We joined the search interaction data with the above SDoH factors on ZIP codes and excluded ZIP codes that did not have either search interactions or census data. The resulting 55 billion search interactions covered web search traffic from 25,150 ZIP codes in the US, and these ZIP codes represent 97.2% of the total US population.

Table B.1 provides per-ZIP code summary statistics of our dataset.

Our choice of search categories was largely informed by our prior work [402]. We chose three determinants – Healthcare Access and Quality, Education Access and Quality, and Economic Stability – from which to draw our search categories. We excluded two determinants that were generally more difficult to capture with simple query string matches because they tend to be more contextual (e.g., location, social) than can be expressed as query strings for information seeking. Under the three SDoH factors, we chose seven search categories that not only appeared more frequently than others in our dataset but also were relevant topics during the pandemic. Table B.4 enumerates the categories we examined with example query strings, URLs, and regular expressions.

Examining individual search keywords or subcategories has been pursued by others within and outside the scope of the pandemic. In our study, the use of broad categories spanning health, economics, education, and food is intended to capture a holistic view of the pandemic across many different needs [347]. Accordingly, we do not make any claims about subcomponents within a category because studying these subcomponents is out of the scope of this work.

Certainly, there exist search keywords that are more popularized by the current pandemic, such as “coronavirus” or “covid”, that also belong in the health information category. However, these keywords are not unique to the current pandemic and have existed before. As infrequent searches for “coronavirus” might seem in 2019, in our data, the query frequency of “coronavirus” in 2019 was similar to that of “mers” and certainly not zero (Figure B.4). In fact, many categories of interests exhibited changes during the pandemic [402, 228], not just some that are highly relevant to the pandemic. For example, Suh et al. [402] have demonstrated that many of the ordinary search topics, such as “toilet paper”, “online games with friends”, or “wedding” were significantly changed during the pandemic.

Disproportional Change in Digital Engagement during the Pandemic

Our goal is to quantify the disproportional change in digital engagement during the pandemic experienced by different subpopulations. Our study conducts several data processing steps and analysis methods to arrive at our findings: (1) we quantify digital engagement by computing relative query proportions for various search categories, (2) we quantify intensification or attenuation of digital engagement by computing changes in digital engagement between before and during the pandemic, and (3) we compare the changes in digital engagement across ZIP code groups.

Digital Engagement Trends We leverage interactions with search engines to obtain signals about digital engagements where everyday needs are expressed or fulfilled through a digital medium, in our case Bing [402]. In our study, we characterize digital engagement through modeling users' search interests as expressions of underlying human needs [402], building upon prior work that uses search interactions to model "interests" which are either expressed explicitly through search queries or implicitly through clicks on results displayed on the search engine result page (SERP) [408, 321, 447]. To gain a nuanced understanding of these search interactions, we categorize each search interaction into topics ranging from health access to economic stability and education access. We match each search interaction to a corresponding category through simple detectors based on regular expressions and basic propositional logic (Table B.4). Each category could have multiple regular expressions applied to either the query string, the clicked URL, or both. Then, we count matching search interactions for a given category. Our query string detectors operate only on English-language keywords such that any cross-cultural or cross-language analysis is out of the scope of this work, but some of our detectors include looking at the click results regardless of the query.

To capture the level of search interest in these categories in relation to all other categories of interest, we compute the proportion of total search queries that belongs to a specific category. For example, we compute the proportion of total search queries that contain health condition keywords such as cancer, diabetes, or coronavirus to quantify the level of interest in engaging in health information-seeking behaviors in relation to all other digital engagement behaviors. In another case, we examine search queries that result in subsequent clicks to state unemployment benefit sites to quantify the level of interest in unemployment benefits.

In addition, the focus on the level of interest through query proportions rather than query frequencies is helpful in our analysis. First, it helps with accounting for the baseline differences in search access between two populations. Second, this focus on relative measures of search query frequency helps adjust for changes in query volume over time, which is a common practice in Information Retrieval and web search log analysis [61, 203]. Figure B.5-Figure B.12 illustrate the temporal variations in relative query frequencies (left) and in relative query proportions (right) in each query category for each of two matched groups across all SDoH factors. Adjusting for the baseline differences in search access allows us to remove the existing access differences between the two groups, and the temporal trends of the query proportions between the two groups become much closer.

Longitudinal Before and During Pandemic Change in Digital Engagement To capture longitudinal changes in search behaviors that are most likely attributable to the pandemic, we use a difference-in-differences (DiD) method [234] to apply several corrections. DiD is often used in econometrics and public health research as a quasi-experimental research method to study causal relationships where a randomized control trial (RCT) is infeasible [452]. Using DiD design with the pandemic as the treatment cannot lead to any causal claims because there is no control group or a counterfactual (i.e., everyone is exposed to the pandemic). In our study, we leverage DiD method to quantify the intensification or attenuation in search behaviors by removing seasonal variations and normalizing on pre-pandemic baselines.

After we categorize each search interaction with our categories of interest, we count and aggregate them per time window (i.e., 2-week or 4-week intervals in our analysis) and per ZIP code (Fig. 3.6a). We compute the proportion of the total query volume represented by each category for these time windows to quantify the level of search interests in that category while removing undesired variations in the query volume over time (Fig. 3.6b). We denote the digital engagement at time t in category c as the fraction of the total number of queries at time t: $E(t, c) = N(t, c)/N(t)$. From this, we control for yearly seasonal variations by subtracting the digital engagements of 2019 from that of 2020: $E(t^{2020}, c) - E(t^{2019}, c)$. People tend to behave differently on weekends, and we observed a 7-day periodicity in our data, sometimes known as the “weekend effect” [353].

Therefore, when comparing two years, it is important to account for the weekend effect. In order to highlight the actual differences that are not explained by weekend mismatches across years, we aligned the day of the week between both years (i.e., Monday, January 6, 2020 is aligned to Monday, January 7, 2019). In addition, we ensured that our comparison analysis included all seven days of the week (i.e., look at means across one or multiples of a full week) (Fig. 3.6c).

Finally, to compute the change in digital engagement during the pandemic since the time at which the US national emergency was declared on March 16, 2020, we subtract the query proportions between January 6, 2020 and February 23, 2020, a period we defined as the "pre-pandemic baseline" (Fig. 3.6d). Even though the national emergency was declared three weeks later, we use February 23, 2020 as the cut-off because individual states declared a state of emergency at different times between February 29 and March 15 of 2020, and to avoid partial weeks in our analysis. Our estimate of the *relative change in digital engagement* in category c between before and during the pandemic is defined as:

$$C(t_{\text{before}}; t_{\text{during}}, c) = \left[E(t^{2020}_{\text{during}}, c) - E(t^{2019}_{\text{during}}, c)\right] - \left[E(t^{2020}_{\text{before}}, c) - E(t^{2019}_{\text{before}}, c)\right] \tag{3.2}$$

Or the relative *percentage change* in digital engagement C_{perc} is expressed as:

$$C_{\text{perc}}(t_{\text{before}}; t_{\text{during}}, c) = \frac{\left[E(t^{2020}_{\text{during}}, c) - E(t^{2019}_{\text{during}}, c)\right] - [E(t^{2020}_{\text{before}}, c) - E(t^{2019}_{\text{before}}, c)]}{[E(t^{2020}_{\text{before}}, c) - E(t^{2019}_{\text{before}}, c)]} \times 100 \tag{3.3}$$

We acknowledge that there may exist a ZIP code with zero or very few search interactions for a given category, especially before the pandemic and in 2019. For example, "stimulus check" may only be relevant during the pandemic for certain ZIP codes. We cannot exclude these ZIP codes because we want a good representation and distribution of ZIP codes in our analysis. If a ZIP code makes only a handful of search queries on various health conditions, for example, but the number of queries increases dramatically due to concerns surrounding comorbidities and health complications, that is precisely the signal we hope to capture and observe across ZIP code groups. We mitigate this potential challenge of zero or near-zero baseline issues in several ways. (1) Our regular expressions are inclusive of potential variations in expressing the categories, including expressions that are likely to occur before the pandemic and in 2019. (2) We aggregate search interactions in two or four-week windows, which consequently reduces the likelihood of having no or very little search interaction before the pandemic. (3) We also aggregate across thousands of ZIP codes that belong to a specific group (e.g., a group of ZIP codes with median household income greater than \$55,224), where the likelihood of having no or very little search interaction before the pandemic for each group is 0%. (4) Instead of computing per-ZIP code DiD, we compute per-group DiD. In other words, we perform a within-group summation before taking the difference, which allows us to characterize the change in digital engagement for a typical ZIP code in the group.

Comparisons across ZIP Code Groups Finally, we aggregate these changes in digital engagements across two comparison ZIP code groups for each SDoH factor, for example, to compare the average change in the digital engagement of low income ZIP codes with the average change of the high income ZIP codes (Fig. 3.6d). Thus, we operationalize the disproportional change in digital engagement during the pandemic by quantifying the differences in the changes in search behaviors for a single search category between two ZIP code groups delineated by a single SDoH factor (Fig. 3.6). In our analysis, we report the change in digital engagement as the percentages of the pre-pandemic baseline, C_{perc}, where 0% denotes no change. We report the disparities in the changes in digital engagement between two comparison ZIP code groups as the percentage point difference where 0 denotes no difference (Fig. 3.6e,f). We formalize disparities in the changes in digital engagement in category c during the pandemic between high-risk ZIP code group g_{high} and low-risk ZIP code group g_{low} as:

$$D_{\text{perc}}(t_{\text{before}}; t_{\text{during}}, g_{\text{low}}; g_{\text{high}}, c) = C_{\text{perc}}^{g_{\text{high}}}(t_{\text{before}}; t_{\text{during}}, c) - C_{\text{perc}}^{g_{\text{low}}}(t_{\text{before}}; t_{\text{during}}, c) \tag{3.4}$$

To obtain non-parametric 95% confidence intervals, we conducted bootstrapping with replacement at 500 iterations during this aggregation step. These confidence intervals are computed when estimating the effect size (i.e., the difference between matched groups) and are visualized with figures demonstrating the difference between groups. All error bars in the figures indicate this 95% bootstrapped confidence interval (N=500). Figure B.13-Figure B.26 illustrate percent changes in each query category for each of the two matched groups and their differences in percentage points across all SDoH factors.

Matched Comparison Groups

Our goal is to quantitatively estimate the independent association between one socioeconomic factor and the changes in digital engagement while controlling for other factors during a global crisis such as the COVID-19 pandemic. Specifically, we are interested in eight SDoH factors: (1) median household incomes, (2) % unemployed, (3) % with health insurance, (4) % with Bachelor's degree or higher degrees, (5) population density, (6) % Black residents, (7) % Hispanic residents, and (8) % with internet access.

One way to do this is to conduct a simple univariate comparison between the two groups. However, one would quickly realize that the high income group has a fewer minority races than the low income group, making the comparison unfair. Many of the socioeconomic and racial variables are known to be correlated [5, 210, 58]. This means that univariate analysis of outcomes along one SDoH factor would likely be confounded by multiple other variables. In fact, within our dataset, we observed a high correlation among many SDoH factors examined (Table B.3). For example, the median household income of the ZIP codes in our dataset is negatively correlated with the percentage of Black residents (Pearson $r = -0.23$) and is positively correlated with internet access (Pearson $r = 0.66$). Comparing high and low income groups without considering other factors would result in two groups of uneven distributions of race and internet access, among many other factors. Therefore, it is important to consider these factors jointly and adequately control for SES factors when analyzing outcome disparities [5, 58]. To create a comparable and balanced set of groups with similar covariate distributions, we leverage matching-based methods.

Matching-based methods are commonly used to replicate randomized experiments as closely as possible in situations when randomized experiments are not possible from observational data [397, 182]. This is achieved by obtaining a balanced distribution of covariates in the treated and control groups [22, 397]. Even though matching-based methods are commonly used for causal inferences, the same matching-based method can also be used to answer noncausal questions [397] (e.g., racial disparities [360]). Our study, therefore, performs a longitudinal before-after observational study with matched groups to answer noncausal questions of the form: “How did the changes in search behaviors during the pandemic differ across matched groups delineated by a single socioeconomic and environmental factor?” In addition, our approach follows best practices for balancing comparison groups in longitudinal studies [259] which we discuss in detail below.

In our study, we apply matching-based methods while considering the SDoH factors as “treatments”. Prior SDoH research suggests that the five SDoH are interrelated and impact one another [377]. Because of this relationship and known correlations between the SDoH factors, we consider all other SDoH factors as potential confounders of a selected treatment factor. It is true that considering SDoH factors as the treatment poses challenges in the framing of the task because these factors are generally not modifiable (e.g., race) or difficult to modify (e.g., income). However, we refer to SDoH factors as treatments, not because they are modifiable, but because we apply the standard formulation of matching-based methods. Identifying modifiable factors in a matching-based experimental study can be used directly to make changes to those treatment factors and to reduce risk. On the other hand, identifying non-modifiable factors has been shown to also be useful to determine high-risk groups that require shielding and targeting for interventions [188].

Because of the high degrees of spatial segregation in the US [143, 449], matching every ZIP code can be challenging. For example, for every ZIP code with low income and high proportions of Black residents, it is difficult to find a unique ZIP code with high income and high proportions of Black residents. Therefore, we perform one-to-one matching of ZIP codes with replacement and achieve better matches (i.e., lower bias). Theoretically, this is at the expense of higher variance, but given the size of our dataset, this downside was not a problem in practice. We use the MatchIt package [187] with the nearest neighbor method and Mahalanobis distance measure to perform the matching.

We leverage an extensive and iterative search across multiple matching methods to achieve maximum covariate balance and representative samples [212]. Regardless of which matching method is superior, one thing to note is that using a “better” matching method does not generally guarantee a better experimental design. It is then common practice to assess the quality of covariate balance, and in the end, it does not matter how this balance was achieved, as long as it was achieved. We choose to perform matching on all covariates, instead of propensity scoring [342] which summarizes all of the covariates into one dimension. Importantly, we demonstrate in Section *Evaluating Quality of Matching Zip Codes* that this method leads to high-quality matches that are balanced across all covariates.

Determining Treatment and Control Groups For each of the SDoH factors, we first split all available ZIP codes into treatment and control groups using a threshold. We use a value close to the median to split the population into two groups for median household income ($55,224), % unemployed (3.0%), % with insurance (92.7%), % with internet access (81.8%), and % with Bachelor's degree or higher (21.1%) because the mean and median of those factors across the ZIP codes are similar. In other cases, the distribution across the ZIP codes is highly skewed. For race/ethnicity, we use the rounded percentage of the national population for that race/ethnicity (12% for Black and 18% for Hispanic residents). For population density, we follow previous practices of urban-rural classification at 500 people per square mile [327]. Table B.1 and Table B.2 outline descriptive statistics of our ZIP codes across SDoH factors as well as the national average and our chosen cutoff thresholds.

We consistently defined the treatment group as "high-risk" according to each of the dimensions of variation we specified [117]. Therefore, our treatment groups are as follows: low income, high percentage of minority residents, low level of educational attainment, high unemployment rate, low insurance rate, low level of internet access, and high population density. For example, for income, we split the ZIP codes into a high-income group (median household income $>$ $55,224) and a low-income group (median household income $\leq$ $55,224), where the low-income group is the treatment group. Then, for each treatment ZIP code, we look for a control (i.e., "low-risk") ZIP code that closely matches it on all other SDoH factors (i.e., $|SMD| < 0.25$ to generate a matching pair of ZIP codes). We performed this matching on all ZIP codes, and we discarded ZIP codes for which we cannot find a good match. As demonstrated in Table B.6, this process retains at least

99.8% of the treatment ZIP codes in our matching process and the discarding of ZIP codes is a rare exception.

Evaluating Quality of Matching Zip Codes To gauge whether two ZIP code groups are similar across the SDoH factors and to determine the quality of matching while minimizing the potential confounding effects of these factors, we leverage Standardized Mean Difference (SMD) across ZIP code groups as our measure of comparative quality. The SMD is used to quantify the degree to which two groups are different and is computed by the difference in means of a variable across two groups divided by the standard deviation of the one group (often, the treated group) [346, 398, 22]. In our analysis, we use $|SMD| < 0.25$ across all our SDoH factors as a criterion to determine that the two groups are comparable, following common practice [397, 398]. For example, when we split our ZIP codes in half along median household income to create a high-income ZIP code group (median household income $>$ $55,224) and a low-income ZIP code group (median household income $\leq$ $55,224) and examine the SMD of other SDoH factors, we find that all SDoH factors except % Hispanic residents and population density fail to achieve the necessary matching criteria of $|SMD| < 0.25$ prior to matching. This means that low-income ZIP codes are more likely to have less internet access, lower educational attainment level, less health insurance, more unemployment, and higher proportions of Black residents. We perform this evaluation process for all comparison groups to find that correlations among all SDoH factors pose threats to validity in univariate analyses. Table B.5 summarizes the mean SMD if we were to directly compare two ZIP code groups created by splitting the ZIP codes along the chosen split boundaries. Instead of such direct comparison, we perform matching and tune the caliper of the matching algorithm to determine a good match and to meet the $|SMD| < 0.25$ criterion between the two comparison groups across all covariates. Table B.6 summarizes the result of the matching operation with the maximum $|SMD|$ being below 0.25, that is ensuring comparability across all covariates, between two ZIP code groups along all SDoH factors. Table B.7-Table B.22 enumerate pre- and post-matching balance assessments between groups for each SDoH factor.

Estimating the Effect Size After identifying treatment and control ZIP code groups with comparable distributions along all SDoH factors, we compare the outcomes (i.e., constructs of digital engagement such as online access to health condition information) between the matched ZIP code groups. This matching process estimates, for example, the differences in the changes in online health information-seeking behaviors between high and low income groups during the pandemic while removing plausible contributions from all other observed factors. The differences estimated in this study help identify high-risk groups (e.g., low income, low educational attainment, high proportions of minority residents) for whom to suggest interventions or targeted shielding to mitigate or reduce risk [188].

It is important to note that our matching process only partially incorporates what Helsper calls the digital impact mediators of access, skills, and attitudes [180]. First, where digital access is concerned, though all search queries in the study presume some form of internet access, we do sample ZIP codes with varying levels of aggregate internet access, allowing us to control to some extent for internet access at the population level. It is important to note, however, that our study lacks the data to account for any changes in ZIP code level internet access during the pandemic due to remote work. Where digital skills are concerned, we do not incorporate direct measures of such technical or operational skills at either the individual or aggregate level, but we do incorporate measures of educational attainment such that we can partially control for this factor in our analysis. Finally, we do not control for individual-level or aggregate-level variation in attitudinal impact mediators such as self-efficacy, as that would be outside the scope of the study. Additional more detailed data would have to be collected and analyzed in order to fully disentangle the impacts of the SDoH factors under study here from such digital impact mediators.

3.2.2 Results

Health information access

First, we examine the proportion of queries relating to a variety of health conditions (e.g., coronavirus and other health conditions including cancer or diabetes). Because the coronavirus, as the underlying cause of the pandemic, is at the forefront of everyone's minds, the relative change in queries related to health conditions is almost 1000% higher than the pre-pandemic baseline. If all things were equal, we would see the same volume of response (i.e., the same relative change in query proportions) across all ZIP codes. However, given the higher rate of pre-existing health conditions, documented disparities in healthcare access, and higher COVID-19 case and mortality rates for low SES subpopulations [5, 88], we would expect to see that ZIP codes characterized by low SES would experience a greater intensification in their need for health information across a variety of health conditions and therefore increase their level of health information seeking behaviors more than their counterpart ZIP code groups. Instead, we find that ZIP codes associated with lower incomes show over a 200 percentage point smaller increase (95% CI $[-287, -152]$) in health condition queries than their higher income counterparts (Fig. 3.6e). This means that a ZIP code that was yielding a thousand health condition queries per month before the pandemic makes about ten thousand such queries per month during the pandemic, but a similar ZIP code would only yield about eight thousand such queries per month if that ZIP code had lower median household income. We find that ZIP codes with higher proportions of Hispanic residents, higher population densities, and higher unemployment rates also responded to the pandemic with a lower relative change in their health condition queries during the first four weeks (Fig. 3.6f). While ZIP codes with high (i.e., above population-average) proportions of Black residents ($\geq$12%) do not seem to be affected as much as those with high proportions of Hispanic residents during the first four weeks, their response is lower during the months of August to November (Figure B.14g). On the other hand, we find that ZIP codes with lower educational attainment ($\leq$21.1% with bachelor's degrees) generate over 70 percentage points more (95% CI $[31, 117]$) health condition queries compared to ZIP codes with higher educational attainment (Fig. 3.6f).

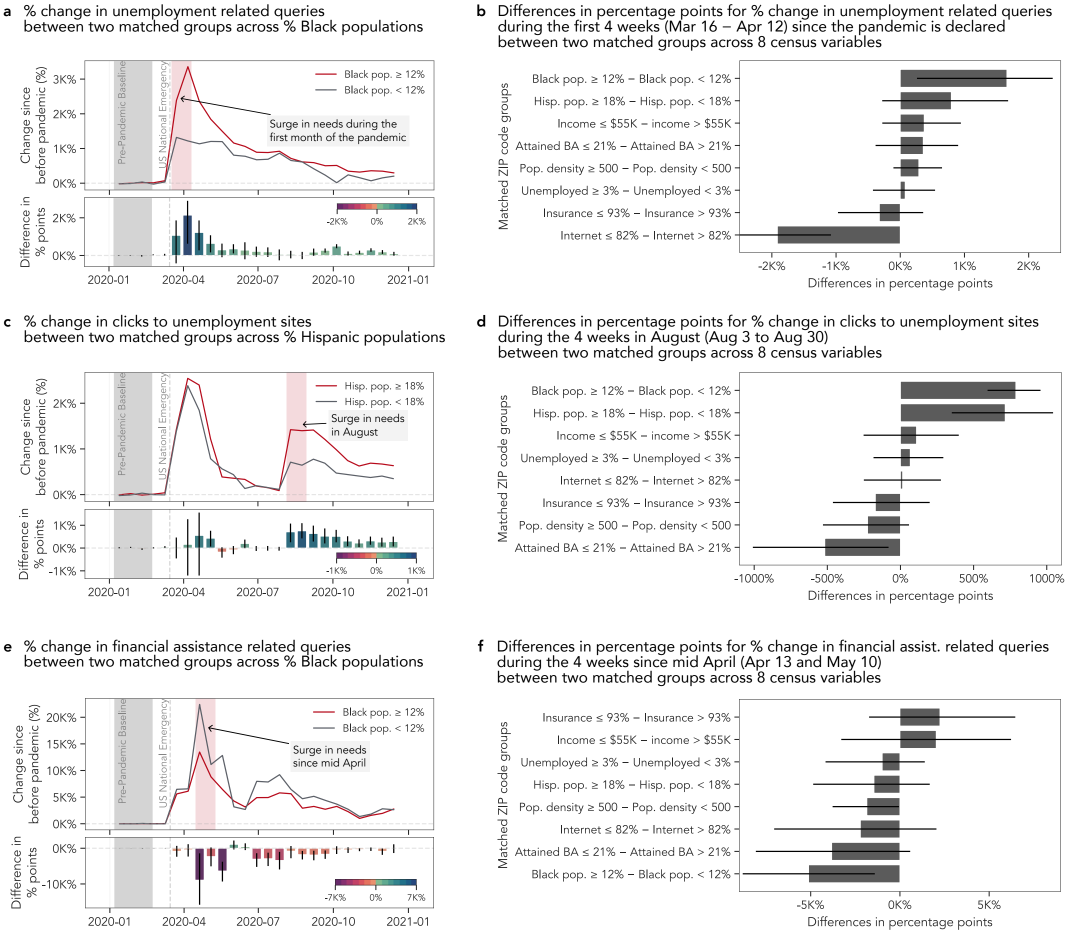
Prior research has shown that SES and demographics correlated with online health information-seeking behaviors, highlighting the digital divide in health information access [100, 155]. This divide has serious consequences. Through effective online health information-seeking behaviors, individuals can potentially make better healthcare choices and enjoy better health and well-being as a result, thereby reducing health disparities [165, 338, 166, 100]. Unfortunately, our results suggest that disadvantages underlying certain socioeconomic contexts of ZIP codes (e.g., income, higher proportions of minority residents) independently are associated with attenuated participation in online health information-seeking behaviors relative to their counterparts. According to prior digital divide research [117, 338, 116], a gap in health information-seeking behaviors may exacerbate health disparities down the line.

Economic assistance access

During economic hardships and especially during the pandemic, the internet can be an efficient way for governments and institutions to deliver interventions and can lower barriers to accessing economic assistance or welfare services (e.g., https://www.usa.gov/food-help provides a comprehensive list of resources for food assistance). Unfortunately, the pandemic imposes multi-layered barriers to accessing crucial economic assistance because low SES subpopulations are more likely to suffer economically from the pandemic [141] and deprioritize improving digital access as a consequence [338]. To understand changes in economic search behaviors during the pandemic, we examine behaviors for accessing unemployment and financial assistance on the web.

When we examine unemployment-related search interactions, we find that relative changes in unemployment-related search queries (e.g., “eligible for unemployment benefits”, “jobless claims”) closely follow those of reported unemployment claims by the Bureau of Labor Statistics (Figure B.2a). However, the intensification of unemployment search queries in ZIP codes with higher proportions of Black residents is almost three times the increase corresponding to ZIP codes with lower proportions of Black residents (Fig. 3.7a), with a 3026% increase in query proportions for ZIP codes with higher proportions of Black residents compared to an over 1365% increase for their counterparts, resulting in a 1,661 percentage point difference (95% CI $[260, 2374]$) (Fig. 3.7b).

Potential interest in digital unemployment resources is not captured in reported claims that measure unemployment claims that are actually submitted, but it can be readily observed in web search logs. For example, we find another surge in search queries that resulted in an over 1000% increase in the proportion of clicks on state-specific unemployment websites past July



a % change in unemployment related queries between two matched groups across % Black populations

b Differences in percentage points for % change in unemployment related queries during the first 4 weeks (Mar 16 – Apr 12) since the pandemic is declared between two matched groups across 8 census variables

c % change in clicks to unemployment sites between two matched groups across % Hispanic populations

d Differences in percentage points for % change in clicks to unemployment sites during the 4 weeks in August (Aug 3 to Aug 30) between two matched groups across 8 census variables

e % change in financial assistance related queries between two matched groups across % Black populations

f Differences in percentage points for % change in financial assist. related queries during the 4 weeks since mid April (Apr 13 and May 10) between two matched groups across 8 census variables

Figure 3.7: Disparities in online economic assistance access. (a) During this first month, ZIP codes with higher proportions of Black residents have expressed up to 3,358% more unemployment related queries. (b) ZIP codes with higher proportions of Black or Hispanic residents, and lower income populations experience greater changes in unemployment related queries during this first month. (c) There is a second surge in search queries that led to clicks in state unemployment sites in August, with ZIP codes with higher proportions of Hispanic residents experiencing more than double the change in clicks in state unemployment sites. (d) ZIP codes with higher proportions of Black and Hispanic residents experience greater change in clicks in unemployment sites during the month of August. (e) Search queries related to financial stimulus were at their peak in late April, right after the time that the first stimulus checks were deposited on April 11. (f) However, ZIP codes with higher proportions of Black residents experienced a smaller change in financial stimulus related queries than ZIP codes with lower proportions of Black residents. In bar charts, data are presented as mean values, and error bars indicate 95% confidence intervals via bootstrapping (n=500).

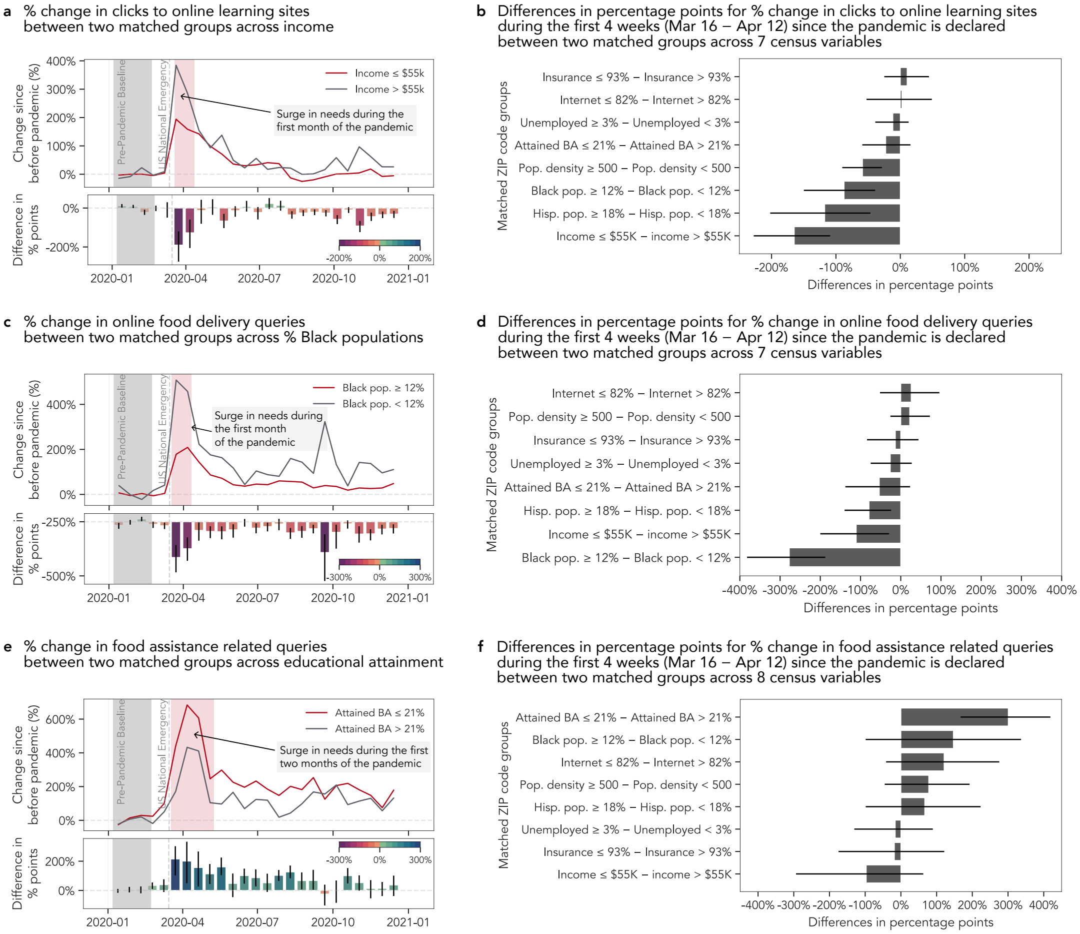
2020 (Figure B.2b), at which point the expanded federal supplement to unemployment insurance benefits expired (Fig. 3.7c). During the month of August, ZIP codes with higher proportions of Black and Hispanic residents present 789 (95% CI $[595, 957]$) and 716 (95% CI $[351, 1043]$) percentage points more in their change in clicks to unemployment sites, indicating that ZIP codes with higher proportions of Black and Hispanic residents may have required additional long-term unemployment benefits. Conversely, ZIP codes with lower educational attainment levels experienced 517 percentage points less (95% CI $[-1009, -81]$) in the change in state unemployment site visits (Fig. 3.7d). Such discrepancy between interests in unemployment benefits expressed online and officially submitted claims and the relatively attenuated access to such resources may suggest potential barriers in the successful submission of benefit applications (e.g., confusion, eligibility [367, 46]). Coupled with a low recipiency rate of unemployment benefits [276] and the association between unemployment accessibility and suicide risks [211], the mismatch between demands and claims is concerning.

April of 2020 was a prime occasion for financial assistance related queries (e.g., “loan forgiveness”, “stimulus check deposit”) because the first stimulus checks were deposited on April 11, 2020 (Fig. 3.7e). We find that financial assistance related queries increased by over 15,000% in mid-April on average, but ZIP codes with higher proportions of Black residents experience 5,119 percentage points less change (95% CI $[-8809, -1407]$) in financial assistance related queries between April 13 and May 10, 2020 (Fig. 3.7f). That means that if a ZIP code yielded 100 financial assistance related queries per month in mid-April of 2019, that ZIP code yields 16,700 such queries per month in mid-April during the pandemic, but only 11,600 queries for an otherwise similar ZIP code with a higher proportion of Black residents. Since we successfully controlled for other potential confounding factors such as income and educational attainment in our comparison, as shown in Table B.8, our result points to higher proportions of minority residents within ZIP codes, not necessarily the racial composition of the ZIP codes per se and certainly not the race/ethnicity itself, as a plausible source for such disparity. Our finding highlights the need to further investigate potential barriers or disadvantages unobserved in our data that disproportionately prevent ZIP codes with higher proportions of Black residents from responding to pandemic-induced stimulus demands on the web.

Shift to digital learning and food delivery resources

The COVID-19 pandemic brought a rapid and massive digital transformation to lives as mandated lockdowns forced people to transform and reimagine traditional interpersonal connections (e.g., going to school, getting food, or meeting friends) into virtual or digital ones. Unfortunately, digital inequalities worsen social and material deprivations and perpetuate existing disadvantages into a "digital vicious cycle" [39, 34]. To observe changes in education search behaviors during the pandemic that may be useful to understand this vicious cycle, we investigate two types of digitally mediated activities that would be presumed to be particularly sensitive to pandemic-induced limitations on in-person access: online remote learning and online food delivery services.

Statewide mandates in the US required many schools to close in-person learning as early as March 16, 2020 [323], and school districts scrambled to implement remote learning alternatives. Many parents, students, and teachers turned to free online resources such as Khan Academy to fill the gaps temporarily or permanently [308]. There were also reported disparities in access to technologies or live virtual learning as well as absenteeism that stymied low income students [183]. When we examined search queries that result in visits to free online learning resources (e.g., coursera.org, khanacademy.org), during the first four weeks of the pandemic, there was an overall increase in the proportion of queries that led to online learning sites compared to before (seen as a positive percent change in Figure B.21). During this time, we found that ZIP codes with lower income and higher proportions of Hispanic residents exhibited only half to two-thirds of the increase (percentage point difference 95% CI $[-227, -109]$ and $[-202, -46]$ respectively) in those queries relative to their counterpart groups (Fig. 3.8a). If a ZIP code yielded 100 search-led clicks to online learning sites per month before the pandemic, that same ZIP code would yield 500 such clicks per month during the pandemic, but only 300 such clicks would be observed for a similar ZIP code with lower income or a higher proportion of Hispanic residents, even after controlling for internet access (Fig. 3.8b). ZIP codes with higher proportions of Black residents and higher population densities exhibit a similar trend. Even though these free online learning



a % change in clicks to online learning sites between two matched groups across income

b Differences in percentage points for % change in clicks to online learning sites during the first 4 weeks (Mar 16 – Apr 12) since the pandemic is declared between two matched groups across 7 census variables

c % change in online food delivery queries between two matched groups across % Black populations

d Differences in percentage points for % change in online food delivery queries during the first 4 weeks (Mar 16 – Apr 12) since the pandemic is declared between two matched groups across 7 census variables

e % change in food assistance related queries between two matched groups across educational attainment

f Differences in percentage points for % change in food assistance related queries during the first 4 weeks (Mar 16 – Apr 12) since the pandemic is declared between two matched groups across 8 census variables

Figure 3.8: Disparities in shifting to digital resources. (a) Online learning sites played a significant role in filling in the gaps introduced by school closures at the beginning of the pandemic with an over 200% increase in engagement. (b) ZIP codes with lower income and higher proportions of Black residents tend to access online learning resources less. (c) With mandated lockdowns, populations have transitioned to food delivery services during the pandemic, but the rate of change in online food delivery queries is more than twice for ZIP codes with lower proportions of Black residents. (d) We see that ZIP codes with higher proportions of Black or Hispanic residents or lower income experienced a smaller change in online food delivery services during the first month of the pandemic. (e) Food assistance related queries were also in high demand with an over 400% increase at the beginning of the pandemic. (f) ZIP codes with lower educational attainment experienced a greater change in food assistance related queries than ZIP codes with higher educational attainment. In bar charts, data are presented as mean values, and error bars indicate 95% confidence intervals via bootstrapping (n=500).

resources are designed to be accessible and flexible, helping students to go at their own pace, we find that ZIP codes with low income or high proportions of Black or Hispanic residents did not leverage them at the same level as their counterpart ZIP code groups during the pandemic.

On the other hand, during the fall academic period of 2020, the proportion of queries that led to online learning sites decreased compared to before (seen as a negative percent change in Figure B.21). During this time, we found that ZIP codes with lower income and higher unemployment rates exhibited a smaller attenuation (i.e., their change remained closer to the baseline, Figure B.22e and h), but ZIP codes with a higher proportion of Black residents exhibited a larger attenuation (Figure B.22g).

In addition, school districts in low SES neighborhoods were more likely to be closed during the pandemic and less equipped to provide remote learning or at-home assignments, greatly reducing opportunities for both in-person and online learning for students with negative educational outcomes [306, 132]. Our findings suggest that there exists unintended consequences of the public health policies that perpetuate a myriad of disadvantages, as education is such a crucial factor in digital literacy [143, 339], income [160], and health [343, 338].

COVID-19 fundamentally changed people's purchasing and spending behaviors, as many of the restaurants, stores, and non-essential businesses were closed to in-person shopping [24]. Spending on food delivery and groceries also increased significantly during the pandemic, with more people eating at home with a higher utilization of online e-commerce platforms for accessing food and groceries [24, 80]. When we examine search queries for online food delivery (e.g., "grocery delivery", "deliver food"), we find that online food delivery queries increased by over 500% for ZIP codes with lower proportions of Black residents while those with higher proportions of Black residents only increased by over 170% (percentage point difference 95% CI $[-382, -188]$, Fig. 3.8c,d). We found similar lessened engagement in online food delivery searches for ZIP codes with lower income and higher proportions of Hispanic residents (95% CI $[-200, -29]$ and $[-140, -24]$ respectively, Fig. 3.8d). These findings could be explained by the fact that low income subpopulations receive and seek more food assistance and tend to eat food away from home less frequently [351] and that such online food delivery services may not be accessible because they incur higher costs for consumers, given the markup and delivery surcharges.

ZIP codes with lower educational attainment also experienced a 301 percentage point higher increase (95% CI $[167, 419]$) in queries for seeking food assistance (e.g., "Supplemental Nutrition Assistance Program", "help with food stamps", "free and reduced lunch", Fig. 3.7e,f) relative to their highly educated counterparts. Unfortunately, those that relied on these traditional food assistance programs were left with severely limited choices during the pandemic because these programs do not extend to online purchase or delivery services [66]. Our findings highlight a potential gap between the increased food assistance need, as illustrated by the increase in the online information-seeking behavior *about* food assistance, and the ability to actually procure food goods through online food purchase and delivery services.

3.2.3 Discussion

We conducted a longitudinal study during the pandemic to observe the second-level digital divide at population scales. Specifically, we leveraged the centrality of web search engines for online information access to quantify how offline exclusion relates to the intensification and attenuation of existing digital exclusion during the pandemic. Our use of search data provided a near real-time and unique lens into naturalistic digital behaviors [127]. Our analysis revealed potentially unmet needs that are unobserved by other data sources. For example, we observed a surge in unemployment site visits in August 2020 that are not captured by the unemployment claims data. We also observed differential uses of pandemic-relevant online resources that span across health information, learning, and food delivery. Most importantly, we demonstrated a disproportionate change in a community's use of these digital resources across several socioeconomic and environmental offline factors. These differences are significant when put into the context of the bidirectional nature of digital and offline exclusion where the lack of the ability to capitalize on digital resources could lead to negative downstream offline outcomes [180, 74, 322].

Our study is structured around the SDoH, a framework commonly used and well-known in public health and disparities research, not only to cover a broad spectrum of factors but also to identify opportunities that promote future research around specific determinants. Under the Economic Stability determinant, we found that the lack of economic stability (median household income, % unemployment) is associated with a smaller increase in engagement in health information-seeking behaviors or online learning consumption at the onset of the pandemic, compared to their higher income counterparts. Under the Social and Community Context determinant, we found that ZIP codes with higher proportions of minority residents (% Black residents, % Hispanic residents) exhibited smaller increases in health information seeking, online learning, or online food delivery behaviors, indicating that these groups fell behind in the digital shift catalyzed by the pandemic [39]. Unemployment-related queries were increased the most by ZIP codes with higher proportions of Black residents at the onset of the pandemic. Unemployment-related site visits were increased the most by ZIP codes with higher proportions of Hispanic residents beyond August, indicating the second wave of potentially unmet demand for unemployment assistance. Under the Education Access and Quality determinant, we found that lower educational attainment (% with BA or higher) is associated with a larger increase in health information-seeking and food assistance-seeking behaviors. Under the Neighborhood and Built Environment determinant, we found that higher population density is associated with a smaller increase in health information-seeking and online learning behaviors. Although internet access was not a variable we examined through matching, per our focus on the second-level divide, we found that the lack of internet access does associate with lower unemployment queries. Because we controlled for all other SDoH factors when comparing groups delineated by a single SDoH factor, our findings have implications for designing determinant-specific interventions and also for examining their potential long-term impacts. Although there are factors we did not find to be significant (e.g., % with healthcare), we caution against interpretations of such factors or interventions not being useful or necessary.

Our analysis along the SDoH factors probes into plausible sources of disproportionate digital behaviors only at ZIP code levels, and understanding the disadvantages underlying these factors and mechanisms for such disparities that permeate through the life course of an individual must be further investigated. In recent years, the SDoH has been referenced in relation to the digital divide; digital literacy and internet access are referred to as “super determinants of health” as they relate to all social determinants of health [373]. Recent digital divide literature also raises an alarm for the “third digital divide” (i.e., the differential offline outcomes that people obtain from their use of digital technologies) and highlights the important interplay between different levels of the divide as well as the role of digital capital in bridging online and offline realms [74, 322]. Therefore, our findings frame important research questions on the downstream real-world implications of differential information search behaviors. For example, high priorities must be assigned to understanding the long-term offline impacts of low income communities not leveraging as many online learning resources or communities with higher proportions of Hispanic residents having intensified unmet demands for online unemployment assistance compared to their counterpart groups. Although the SDoH factors and outcomes reviewed in our analysis are generally not modifiable (e.g., race) or difficult to modify (e.g., income), our findings nevertheless highlight specific at-risk populations for whom to target shielding or interventions [188].

Prior studies have shown that access to digital resources and information and the incorporation of such digital technologies in everyday lives from childhood are crucial for upwards mobility [338]. Although SES is an important factor in shaping disparities in digital access, prior research has shown that SES also impacts levels of web expertise and the utilization of digital resources for information-seeking activities [173]. Low SES populations suffer from the lack of training and educational support key to building the necessary skills to make efficient use of digital access and tools [143], highlighting that simply making the internet more accessible may not level the playing field [324]. In the context of the current COVID-19 pandemic, where digital access and resources became more critical due to prolonged at-home isolation and restrictions on in-person activities, communities characterized by low SES may experience the compounding effects of multiple potential disadvantages that may manifest as disparate reactions to the pandemic in digital engagement.

We note the inherent limitations of studying digital engagement using digitally obtained data: This and other studies with online data can inadvertently exclude those who leave no or very little digital footprint [338]. Our information sources provide signals about levels of activity, but we cannot study the details of changes in types of access if there is no engagement. Our analysis is also limited to the footprint of Bing as one of several search engines used for online information access, and Bing's user population may not be fully representative of the United States population. We use both English regular expressions as well as language-independent click-based measures but did not include regular expressions in other languages. Our study carefully controls for internet access, as measured by the census, such that any observed effects cannot be explained by differences in internet access across ZIP code groups. Our observed changes can only be attributed to ZIP code levels and not individuals to preserve anonymity and because individual-level SDoH factors are not available. Our work provides a holistic characterization of digital engagement using broad categories spanning health, economics, education, and food, and we cannot make claims about specific subcomponents (e.g., individual keywords). Our longitudinal comparison between before and during the pandemic cannot be used to isolate the changes in search behavior to be solely attributable to the pandemic to make any causal claims, despite our adjustments for temporal variations.

Our current data cannot be directly used to discern whether different access behaviors are due to the lack of web expertise (i.e., digital literacy or search facility), the lack of awareness of the value of information (i.e., attitude towards information), or the lack of intangible resources like time and energy. However, concepts like digital literacy, which is an important factor in the embodiment of digital capital, can be quantified by careful examination of an individual's search behavior. As prior research has shown, search interactions vary, based on the user's familiarity with search engines or their domain expertise [448, 446]. Quantifying digital literacy combined with a longitudinal observation of socioeconomic and environmental factors could provide empirical evidence for how digital literacy operates in the attainment of offline economic, cultural, and social capitals [74], and our large-scale, search-based methodology opens the doors of opportunities for monitoring such phenomena. In addition, we see value in follow-up, small-scale focused studies aimed at contextualizing individuals' experiences of the crisis and measuring the effects of community-specific interventions [39]. These community-specific interventions could include raising the level of digital literacy (e.g., education around web expertise or digital know-how) or improving the quality of digital access (e.g., high-speed, uninterrupted internet access or high-end equipment). Quality of access, especially through different device types or device specifications, has been highlighted as another important factor in recent digital divide research [423]. Therefore, more work is needed to understand the differential uses on desktop or mobile devices. These may also include non-digital methods because traditional methods (e.g., text messaging, handouts) have been shown to work better for low SES populations [183]. Future research aimed at understanding digital disparities, therefore, must acknowledge the correlations between different SES, race/ethnicity, and social determinants of health [58] and leverage methods that embrace their interrelatedness [451].

This study presents a web-based approach to understanding digital disparities. It demonstrates that web search logs can be harnessed to characterize and deliver key insights about the disproportional utilization of digital resources to meet everyday needs during global crises. Our observational study design is able to scale to a large population (billions of queries by millions of people) to quantify the disparities in digital engagement. Building on prior disparities research that advocated for a comprehensive look at SES factors including race/ethnicity [5, 58], our study emphasizes the inclusion of a broad set of factors and outcomes representative of the SDoH. Through the lens of SDoH factors, our findings highlight disadvantaged communities that may be struggling to overcome burdens induced by the pandemic and have disproportionately intensified or reduced their access to critical online resources. Therefore, future public health interventions should target both potential barriers to access that pull communities away from necessary digital resources as well as provide support to ensure that the intensified need for digital resources is adequately met.

3.3 SUMMARY OF CONTRIBUTIONS TO THESIS

The COVID-19 pandemic has disrupted the lives of people across the world, and much prior research has shown that socioeconomically and environmentally disadvantaged subpopulations have been disproportionately and negatively affected by the pandemic. In this chapter, we aimed to understand the impact of the pandemic in the digital realm, specifically in online information search domains, and to characterize the changes in human needs across populations of different socioeconomic and environmental backgrounds.

We developed a computational methodology that can capture a holistic view of relative changes in online information-seeking behaviors following the pandemic through a difference-in-differences approach. We applied this method to over 35 billion search interactions spanning over 36,000 ZIP codes over a period of 14 months to characterize changes in human needs across physiological, socioeconomic, and psychological domains in the US. Then we investigated how these changes were experienced differently by different communities by analyzing 55 billion search interactions across 25,150 US ZIP codes through a matching-based approach that disentangles the potential confounding effects of socioeconomic and environmental factors.

Our findings showed that the expression of basic human needs increased exponentially while higher-level aspirations declined during the pandemic in comparison to the pre-pandemic period. In exploring the timing and variations in statewide policies, we found that the durations of shelter-in-place mandates influenced social and emotional needs significantly. Across several information domains, including health and pandemic-relevant online resources (e.g., online learning, online food delivery), our findings revealed that there had been a disproportionate change in a community's use of digital resources to address their needs during the pandemic, even if they were digitally connected. Our analysis also revealed potential barriers to meeting unemployment needs that are not captured by the unemployment claims data.

Our work presented in this chapter suggests that signals from web search logs can be used to characterize and monitor human needs and digital disparities at population scales. The major shifts in a spectrum of human needs observed echos the need to understand the societal, economic, and psychosocial effects of these events and the underlying system of human needs. The differences we observed across socioeconomic and environmental factors are significant when put into the context of the bidirectional nature of digital and offline exclusion. Our findings highlighted disadvantaged communities that may need additional or targeted shielding through public health policies and programs that can minimize negative downstream effects.

My work in the pandemic context demonstrated that *computational methods* (e.g., difference-in-differences, matching-based comparisons, longitudinal before-after observational method) combined with holistic frameworks (e.g., human needs framework, social determinants of health) can be leveraged to understand *human needs and disparities at population scales*. Although this work does not implement public health interventions, which are difficult to achieve at population scales, it highlights opportunities for research and development of *community-specific interventions* that can lead to *a higher engagement in digital resources necessary for well-being*.

Chapter 4

SUPPORTING COLLABORATIVE CARE FOR PATIENTS WITH CANCER AND DEPRESSION

In the previous chapter, I situated my research in the context of online search behaviors during the pandemic. Although I demonstrated that using computational methods can surface human needs and disparities at population scales, I did not design public health policies and interventions that would address those needs or disparities. In this chapter, I situate my research in the context of depression care for patients with cancer and exercise all three stages of the human-centered design process I described in Section 1.2, from understanding the stakeholder needs, designing interventions, to deploying systems for evaluation.

Cancer is one of the leading causes of death worldwide, with an estimated 1.7 million people with a cancer diagnosis in the United States in 2018 [436]. Although depression is one of the most common challenges for patients during and after cancer treatments [221], it is severely under-treated in patients with cancer, including about 25% of patients with cancer and depression that receive no depression treatment at all [430]. Because of the complex interactions between depression and cancer treatments [317], directly integrating psychosocial services in cancer settings is highly recommended [137], and care models such as Collaborative Care Management (CoCM) have been proven to be effective in treating depression and to be cost-effective in randomized control trials [21, 410, 78, 126, 197]. However, prior implementations of collaborative care have been met with a myriad of challenges, including low fidelity in adhering to principles of effective collaborative care [425], lack of clarity in role definitions and boundaries, challenges in long-term sustainability, and lack of standard care and communication pathways and tools [370, 453, 286]. In addition, patients with co-morbid cancer and depression face additional challenges due to their weakened physical conditions, competing and overwhelming numbers of appointments, unmet logistical needs (e.g., transportation), stigma, or lack of knowledge around depression treatments. Therefore, it is important to uncover challenges in the specific context of co-morbid cancer and depression to inform the design of technology solutions.

In this chapter, I focus on the clinical context of patients with co-morbid cancer and depression where we design technology-enhanced collaborative care for improving patients' psychosocial outcomes. In one study, I demonstrate that *human-centered methods* applied at *personal and organizational scales* can reveal *multistakeholder challenges* that arise at the intersection of cancer and psychosocial care journeys. To address those challenges through technology support, I *digitize and adapt core components* of evidence-based care strategies in a series of iterative design evaluation and pilot studies and demonstrate that such adaptation can *improve engagement in depression care.*

The first study [400] examines contextual factors to identify challenges and needs in the depression care of patients with co-morbid cancer and depression (Section 4.1). We conducted interviews with 29 stakeholders regarding the care of patients with co-morbid cancer and depression: patients, center administrators, medical providers (i.e., oncologists, psychiatrists), and behavioral health providers (BHPs; i.e., social workers, psychologists) and contextual inquiries with 8 BHPs. Interviews and contextual inquiries were conducted across three urban or rural cancer centers that have implemented or aspire to implement collaborative care (i.e., have implemented some components of collaborative care) (Section 4.1.1). Our findings revealed that patients that navigate both cancer care and depression care face additional barriers to their depression care due to the interactions and sometimes competition between the two care journeys and the cancer-centric processes of care (Section 4.1.2, Section 4.1.3). Inspired by the discovered challenges, this study prioritized technology design opportunities for supporting whole-person care for patients with co-morbid cancer and depression (Section 4.1.4, Section 4.1.5).

Following the technology design opportunities from the formative study, we designed technology enhancements to existing care strategies to improve stakeholder collaboration and engagement in depression care (Section 4.2). We digitized core components of collaborative care and behavioral activation (Section 4.2.1). Using a model for a technology-enhanced collaborative behavioral activation loop, we designed and developed a collaborative behavioral activation system, named SCOPE (Supporting Collaborative Care to Optimize Psychosocial Engagement) (Section 4.2.2). We evaluated it through four rounds of interview studies with patients and BHP stakeholders (Section 4.2.3). Our design studies confirmed that the system does support the collaborative behavioral activation loop, but they also uncovered technical and non-technical opportunities for improvement (Section 4.2.4). The studies also identified future research opportunities in deploying a technology-enhanced collaborative care system in a real-world, clinical context (Section 4.2.5).

The key takeaways and contributions to my thesis from these studies are summarized in Section 4.3. The first study was conducted in collaboration with Amy M. Bauer, Jesse R. Fann, James Fogarty, Emily Friedman, Gary Hsieh, Alejandra Lopez, and Spencer Williams. I contributed to the design of the first study, the data collection and analysis, and the interpretation of results. The first study was published in CSCW 2020 [400] and presented at AMIA Clinical Informatics Conference 2020 [401] and CHI 2020 workshop on Rethinking Resources for Mental Health [399]. The subsequent studies were conducted in collaboration with Lydia Andris, Amy M. Bauer, Jesse R. Fann, James Fogarty, Gary Hsieh, Tae Jones, Ravi Karkar, Ty W. Lostutter, Anant Mittal, Spencer Williams, and Ben Zheng. I contributed to the conceptualization of the iterative design and pilot interview studies, the design and development of the system, the data collection and analysis, and the interpretation of results. The research work from the subsequent studies is yet to be published.

4.1 PARALLEL JOURNEYS FRAMEWORK AND COMPETING DEMANDS

Cancer care has increasingly moved toward integrating psychosocial care directly into cancer services [137], and care models such as Collaborative Care Management (CoCM) have been proven to be effective in treating depression and to be cost-effective in randomized control trials [21, 410, 78, 126, 197]. Collaborative care is a team-based care approach with an oncologist, a care manager, and a psychiatric consultant working together to improve the patient's psychosocial health outcomes through the use of evidence-based care practices and a patient registry to deliver measurement-based and population-based care. Despite the effectiveness of such models, a recent national survey found that approximately 80% of cancer centers lack processes to follow up and adjust treatments for patients with clinical depression [466]. Implementing collaborative care is also not without its own challenges. For example, recent reviews and analyses have highlighted: (1) low fidelity in adhering to principles of effective collaborative care [425], (2) lack of clarity in role definitions and boundaries, (3) challenges in long-term sustainability, and (4) lack of standard care and communication pathways and tools [370, 453, 286]. 20% to 50% of patients are known to prematurely drop out of therapy [405, 375, 374], and patients with cancer and depression face additional challenges to engaging in care due to factors that range from weakened physical conditions, competing and overwhelming numbers of appointments, unmet logistical needs (e.g., transportation), and stigma or lack of knowledge around depression treatments. Behavioral health providers (BHPs) in cancer centers, typically trained in clinical social work, are particularly overburdened with attempting to provide depression treatments in addition to helping patients meet their navigational and financial needs.

Our research goals are: (1) to understand contextual factors and to identify challenges and needs in the depression care of patients with co-morbid cancer and depression, and (2) to identify technology design opportunities to support such care and to enhance or facilitate collaborative care. We took a human-centered design approach to examine existing challenges to collaborative care by conducting interviews and contextual inquiries across three urban or rural cancer centers that have implemented or aspire to implement collaborative care (i.e., have implemented some components of collaborative care). Prior research has examined the use of technology to support collaborative care (e.g., telehealth, electronic health records, web, and mobile platforms), finding it to be effective in engaging patients and treating depression [224, 145, 341, 390]. Although there have been a variety of opportunities proposed for technology enhancements of collaborative care [28, 326], such technologies must be carefully grounded in concrete needs identified using human-centered design approaches [29, 30]. Furthermore, a better understanding of the needs for and feasibility of such technologies in collaborative care, especially for treating depression in cancer settings, is necessary to improve the effectiveness and usability of such technologies [325]. Our research thus augments this growing body of work by uncovering challenges and opportunities for design in the specific context of co-morbid cancer and depression.

Although we observed that patients with co-morbid cancer and depression experience both a cancer care journey and a psychosocial care journey consistent with prior research, our analysis revealed that considering these care journeys independently was insufficient for characterizing these patients and the challenges encountered by them and their care teams. Our research therefore also proposes a framework of parallel journeys, examining challenges and opportunities for interventions and technologies through a lens of placing these journeys side-by-side and examining where they fall out of sync.

Our specific contributions therefore include:

- We propose the concept of a *parallel journeys framework* for examining and characterizing complex co-morbidities and care contexts.
- We conduct interviews with 29 stakeholders regarding the care of patients with co-morbid cancer and depression: patients, center administrators, medical providers (i.e., oncologists, psychiatrists), and BHPs (i.e., social workers, psychologists) and contextual inquiries with 8 BHPs. Interviews and contextual inquiries were conducted at three different cancer centers.
- We confirm that patients with co-morbid cancer and depression experience various phases of a *cancer care journey* as described by Jacobs et al. [198].
- We also observe that patients with cancer and depression experience various phases of a *psychosocial care journey*, a concept described in psychosocial care practice guidelines that we strengthen with qualitative evidence.
- We apply the *parallel journeys framework* to analyzing our data and examining breakdowns and challenges to care that occur at the intersection of the parallel care journeys.
- We present technology design opportunities for addressing these challenges and supporting whole-person care for patients with co-morbid cancer and depression.

4.1.1 Method

Two research questions that guided our study are:

RQ1. What are the challenges and needs that patients with co-morbid cancer and depression and their care team encounter in providing depression care?

RQ2. What are technology design opportunities for supporting patients with co-morbid cancer and depression and enhancing integrated depression care?

To answer these research questions, our study included a combination of semi-structured interviews and contextual inquiries. We conducted interviews with multiple stakeholders to obtain multiple perspectives on current practices and challenges surrounding the care of patients with co-morbid cancer and depression. We also conducted contextual inquiries to complement and contextualize our interview data by observing care practices as they unfold in real-world situations. Although our ultimate goal is to identify technology design opportunities, we ensured that our data collection was not exclusive to technology use in order to explore contextual factors and challenges that may influence the design and adoption of technology.

To ensure the safety of patients and maintain high ethical standards for the recruitment and study procedure, we worked closely with the coordinating institution's Institutional Review Board (IRB) as the single IRB of record, with one behavioral health provider (BHP) from each site, and with psychiatry and behavioral science experts within our research team.

Study Procedure

Sites We conducted our study in three cancer centers that varied across dimensions of: (1) urban to rural according to Rural-Urban Continuum Codes (RUCC) [368], and (2) extent of existing collaborative care implementation. Specifically, sites included two urban (RUCC $= 1$) cancer centers and one rural (RUCC $= 5$) cancer center. One urban site (i.e., Site 1) had implemented collaborative care, including BHPs playing the role of care managers (e.g., clinical oncology social workers) assigned to different cancer specialties, embedded psychiatrists, weekly systematic caseload reviews and consultations between BHPs and psychiatrists, and an Excel-based registry to manage and track patients receiving behavioral health care at the cancer center. The other two sites (i.e., urban Site 2, rural Site 3) were aspiring to implement collaborative care with limited integrated psychosocial care. Both sites offered evidence-based psychosocial treatments through BHPs who were embedded within the cancer center, but they lacked access to dedicated psychiatrists, processes and workflow to support systematic follow-up and measurement-based care, and a registry for caseload management. We chose to study these sites with different characteristics to ensure that our data derived from sites with variation in urban to rural locations and their experience with collaborative care. Table 4.1 (top) summarizes site characteristics.

Due to the sensitive nature of depression and a vulnerable patient population, our study required significant coordination and collaboration from the local sites. We held an orientation meeting with various site representatives from each cancer center (i.e., site administrators, providers, and medical directors) to present our study objectives and general procedures. We then identified primary points of contact (i.e., one BHP from each site) who helped finalize participant recruitment procedures, facilities use, scheduling, and compensation.

Participants At each site, we recruited four types of participants: behavioral health providers (BHP), medical providers (MP), administrators (A), and patients (Pt). BHPs are clinicians responsible for delivering psychosocial treatments and coordinating psychosocial care. These participants primarily consisted of clinical oncology social workers, with the exception of one psychologist. Medical providers consisted of oncologists as well as psychiatrists and primary care physicians. Administrators consisted of a cancer center manager, a clinical services manager, and a clinical social work manager who assumed the role of overseeing the operations and workflow of behavioral care at the cancer center. We relied on points of contact from each cancer center to identify and recruit BHPs, medical providers, and administrators.

Once identified, BHPs helped with patient recruitment in two ways. After providing BHPs with a study information handout and recruitment guidelines, we asked each BHP to identify several of their active patients whose recent PHQ-9 [223] (or equivalent) scores were 10 or above (i.e., moderate to severe depression) and to ask patient permission for the study team to contact them to participate in the interview study. In addition, each BHP contacted and obtained permission from patients whose sessions with the BHP would be observed as part of a contextual inquiry.

Across three sites, we conducted semi-structured interviews with 29 participants (11 patients, 9 BHPs, 6 medical providers, and 3 administrators). Of 11 patients interviewed, 7 reported as female, 3 reported as male, and 1 did not report any gender. Six patients reported as an ethnic

Table 4.1: Site characteristics (top) and interview participants from each site (bottom). Participant categories include patients (Pt), administrators (A), behavioral health providers (BHP), and medical providers (MP).

		Site 1	**Site 2**	**Site 3**
Site Characteristics	Location	Urban	Urban	Rural
	CoCM Implemented	Yes	No	No
Interview Participants	Patients (Pt)	3	3	5
	Administrators (A)	1	1	1
	Social workers (BHP)	4	2	2
	Psychologists (BHP)	0	1	0
	Psychiatrists (MP)	1	0	1
	Primary care physician (MP)	0	1	0
	Oncologist (MP)	1	1	1
	Total	10	9	10

minority. Patient ages ranged from 24 to 89 years old with a mean of 47.0 and a standard deviation of 18.0. Table 4.1 (bottom) also summarizes the number of interview participants from each site. Of the 9 BHPs interviewed, 8 were also shadowed and observed through contextual inquiry. As a result of contextual inquiry, we observed 10 unique patient sessions in Site 1, 9 in Site 2, and 7 in Site 3. We did not collect any demographic information for patients observed as part of a contextual inquiry because these patients were not the primary focus of the contextual inquiry. Of the 26 unique patients observed during the contextual inquiry, one was also interviewed.

Interview We conducted semi-structured interviews of multiple stakeholders to collect their perspectives on care experiences in their own words and to triangulate common challenges and needs. The focus of each interview varied slightly based on the type of participant. With patients, our interview questions focused on understanding their experiences surrounding the depression treatment they received, coordination of psychosocial appointments, communication with their providers, and the practices and tools used in addressing their depression and symptoms. With BHPs, we focused on their workflows, tasks, and tools used in providing and coordinating psychosocial care within and outside of their sessions with patients and in communicating and collaborating with different providers. We also probed their experiences in providing evidence-based interventions, specifically behavioral activation, which is an effective treatment for depression [119, 130]. With psychiatrists (or with a primary care physician if the site lacked access to a psychiatrist), interview questions were similar to those for BHPs with added attention to aspects that differentiated their roles (e.g., recommending psychopharmacological interventions or psychiatric consultations). With oncologists, we asked about their involvement in the psychosocial care of their patients, including the detection of depression symptoms, communicating and coordinating care with psychosocial or psychiatric care providers, and managing treatments and medications. With administrators, our interview questions focused on operational and financial aspects of psychosocial care, monitoring the quality of psychosocial care, and adoption of new programs or technologies. These interview topics helped structure our interviews, but our interview protocol was open-ended such that we could further explore challenges in any of these topic areas with all participants.

During the interview, providers were asked to refer to a specific patient encounter to ground their responses (e.g., "please recall a last patient or session that..."). These were not necessarily the same patients we interviewed or observed while shadowing BHPs. To preserve patient confidentiality, we explicitly did not gather patient identifiers from the providers or health records. We instead captured general patient characteristics (e.g., female stem cell transplant patient), behaviors (e.g., checking text messages on the phone), and affect (e.g., anxiety from the infusion) from the conversation. Each interview lasted approximately an hour; 26 were conducted in person and 3 by phone. Interviews were audio-recorded and later transcribed. Each interview participant was compensated with a $50 gift card.

Contextual Inquiry Contextual inquiry [45] allows for observing actual behaviors and situations as they unfold in real-world contexts. The primary focus of our contextual inquiries was BHPs. We were interested in: (1) gaining a richer understanding of existing workflows and practices, competing tasks and demands, and of technology use in the delivery of psychosocial care, and (2) observing a wide variety of patients and treatment techniques employed by the BHPs.

A typical contextual inquiry session started in the morning with a quick orientation to the day's schedule and the patients we would be observing. During observation, we often sat next to the BHPs, listened to their conversations, watched their use of tools (e.g., desktops, mobile devices, printers, notebooks), and traveled with them to their patient sessions and meetings. When observing patient sessions, we did not interfere with the sessions and sat in the room in a way that avoided a patient's direct line of sight. Outside of patient sessions, we asked the BHPs to think aloud if the circumstances allowed it (e.g., charting at their desks), and we asked clarifying questions in the moment or at a later opportune time (e.g., en route to sessions, during lunch breaks).

Each contextual inquiry session with a BHP lasted between 2 to 8 hours and consisted of one or more patient sessions ranging from approximately 15 minutes to an hour. Most patient sessions were scheduled, with the exception of 3 ad-hoc sessions. To address concerns of potentially coercing patients into being observed, each corresponding BHP obtained permission in advance from their patients for our research team to observe their sessions before our visit. We took notes throughout the observation, captured photographs of the facilities and work spaces with the permission of administrators, and received copies of psycho-education materials used in the session. No compensation was offered for participation in contextual inquiries.

Analysis

We combined interview transcripts and observation notes to create a single dataset. All potential identifiers (e.g., names, places) were manually removed by the researchers that collected the data. Our general analysis approach was to draw out challenges and needs for answering RQ1 and then to synthesize technology design opportunities for RQ2. Here we describe our analysis of challenges in greater detail, which involved two stages: (1) enumeration of challenges, and (2) characterization of challenges through patient care journeys.

Identification of Barriers and Facilitators To understand what challenges exist in our data, we identified *barriers* that hinder the progress of depression care and *facilitators* that enable depression care in cancer settings. Three researchers qualitatively coded the interview and observation data using inductive thematic analysis [57]. Two of the three researchers were already familiar with the data, having collected and anonymized most of the data. All interview transcripts and observation notes were entered into qualitative data analysis software (i.e, ATLAS.ti). Researchers independently examined a subset of assigned transcripts and notes, tagging parts of the texts with "memos" [49] that summarize the dialog or observation (e.g., "BHPs are responsible for organizing patient support groups"). We held collaborative memo extraction sessions for a subset of data to align our process and reconcile disagreements. The researchers were encouraged to reuse existing memos if the extracted quote could be described with an existing memo. Although technology-related themes emerged from our analysis, the researchers reframed the technology themes as the underlying stakeholder needs that the technology addresses. We then collaboratively merged, split, and refined extracted memos, using consensus meetings to organize them into emerging barriers and facilitators to psychosocial care. We iteratively repeated these processes of extracting memos independently and collaboratively organizing memos into emerging themes of barriers and facilitators (e.g., "BHPs have many competing demands"). At the end of the coding process, three researchers collaboratively reviewed all barriers and facilitators and the underlying memos and quotes to reach an agreed level of cohesiveness and consistency within each theme. We extracted the underlying challenges that our identified *facilitators* supported and merged those underlying challenges with *barriers* into a combined set of challenges. We incorporated technology-specific facilitators into our consideration for technology design opportunities.

Development of Psychosocial Care Journey After the enumeration of challenges, we organized our data into a temporal care journey to map each challenge to a certain point in time of patient care. Our construction of a psychosocial care journey involved mapping out the timeline of all stakeholder experiences through the care process based on three sources of data: (1) patient recollection of their own care experiences, (2) provider accounts of how their patients experience care, and (3) our observation of patient experiences. On the horizontal, temporal axis, we had various transition points and phases of care informed by psychosocial care practice guidelines [19]. On the vertical, stakeholder axis, we listed the patient and members of the care team and illustrated any actions performed by or interactions between stakeholders. We then

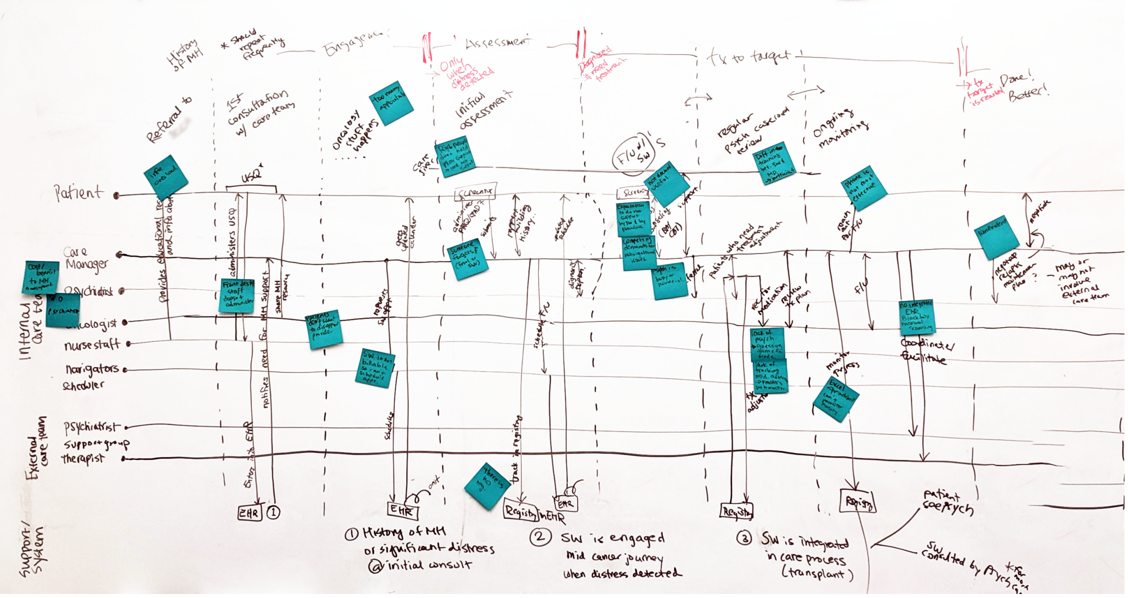


Figure 4.1: A snapshot of the psychosocial care journey mapping exercise for organizing challenges we discovered in our data. The horizontal axis illustrates different phases of psychosocial care, and the vertical axis describes different stakeholders and their interactions.

mapped several challenges at different points of the care journey. Figure 4.1 provides a snapshot of this process. We found this exercise lacked the expressivity required to characterize the complex care paths of patients. This realization led to the development of the *parallel journeys framework* which we describe in detail in Section 4.1.2.

We organize our findings according to our research questions. We first address RQ1, presenting contextual findings that helped us understand the challenges and needs that patients and their care teams face in receiving and providing depression care in cancer settings. Section 4.1.2 describes patient care journeys and our development of the *parallel journeys framework*. Section 4.1.3 applies the framework to characterize challenges identified in our data. We generally found contextual factors to be more prominent in our data because most participants had minimal use of technology in their current care practices. We then address RQ2, presenting technology-specific findings synthesized to address identified challenges. Section 4.1.4 presents technology design opportunities for supporting depression care of patients with cancer and describes why these opportunities are specifically important for our patient and care team context.

4.1.2 Parallel Journeys Framework

Cancer Care Journey

The emotional and physical experiences of patients with cancer can commonly be characterized by phases of a cancer journey [178]: screening and diagnosis, information seeking, acute care and treatment, no evidence of disease, and chronic disease and disease management. Jacobs et al.'s cancer journey framework [199, 198] further characterizes these phases by incorporating patient responsibilities, challenges, and personal impacts in each phase.

The patients that we interviewed, that we observed during contextual inquiries, and that were described by providers, experienced these phases of a *cancer care journey* while managing their patient responsibilities. They described encountering challenges in care and experiencing changes in emotions and life. Pt13, whom we observed while shadowing BHP2, had recently been diagnosed with cancer and was finalizing treatment decisions and financial adjustments in preparation for active treatments; he began feeling a general sense of "*gratitude and compassion*" since the diagnosis. Pt1 described the process of getting her body prepared for stem cell transplant as being "*really rushed,*" and Pt2 explained that, in hindsight, she did not fully process all the diagnostic information given to her because "*people were making [her] feel like everything was not a big deal when it was a huge deal.*" Four of the patients we observed while shadowing BHPs were actively being treated in the infusion suite, and BHP5 described the care environment as "*it's best to always expect things will not go smoothly (laughs).*" When we interviewed Pt1 just after she had finished her transplant procedure, she mentioned that her struggle was in "*trying to figure out what [her] new normal is going to be.*"

Our analysis revealed that the challenges that patients with cancer and depression expressed around depression care could not fit squarely within one or more of these phases of the cancer care journey. For example, psychotropic medications used to treat depression could interfere with cancer treatment, thus requiring significant coordination between the psychiatrist and the oncologist. Pt8 developed severe anxiety towards chemotherapy infusions that required a BHP's physical presence to continue with treatment. Patients often feel too sick from cancer treatments and may cancel their appointments with their BHPs, which impacts the continuity of their depression care. Although the cancer care journey can be used to characterize a range of emotional and psychosocial aspects of cancer (e.g., attitude change, return to normal), its phases are delineated by the physical aspect of cancer (i.e., around the diagnosis, treatment, and recovery from cancer), and it does not account for the context introduced by co-morbid psychiatric disorders that include dedicated treatments and unique patient experiences. Therefore, the cancer care journey alone is insufficient in characterizing such co-morbid contexts.

Psychosocial Care Journey

As described in Section 4.1.1, we generated a timeline of how patients experienced their depression care, organized into four key phases of a *psychosocial care journey*. We found that patients with depression experience four common phases of psychosocial care (see Figure 4.2): (1) identification of patients with depression, (2) initial psychosocial assessment, diagnosis, and rapport building, (3) active depression treatment, and (4) maintenance and relapse prevention planning. These phases of a psychosocial care journey align with practice guidelines recommended for treating major depressive disorder [150], psycho-oncology care practice guidelines [19], and collaborative care guidelines and care responsibilities that include a screening of patients for depressive symptoms, delivering evidence-based behavioral and/or pharmacological treatments to improve patient outcomes, a systematic review of patients with psychiatric consultants and adjustment of treatment for patients who are not improving, and relapse prevention planning [415, 135]. Our data from the patient recollection of care, BHP description of how they engaged with patients, and our observations of BHP sessions with patients at different points of care (e.g., initial engagement, active treatment, follow-up) provide qualitative evidence and a nuanced understanding of how patients with cancer and depression experience these phases. Here we formally introduce the four phases by defining their care goals and providing evidence of how our patients with cancer and depression experience the four phases of the psychosocial care journey.

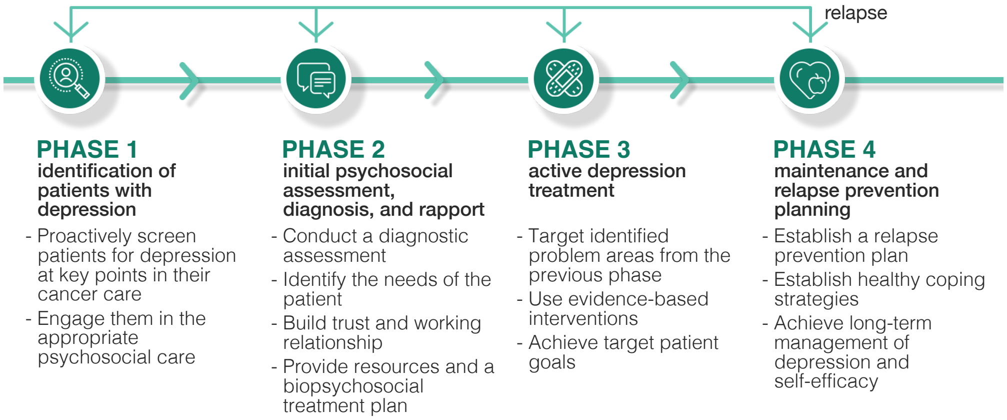


Figure 4.2: Four phases of a psychosocial care journey: (1) identification of patients with depression, (2) initial psychosocial assessment, diagnosis, and rapport building, (3) active depression treatment, and (4) maintenance and relapse prevention planning. Patients may experience a relapse and fall back to the first three phases.

Phase 1: In the **identification of patients with depression** phase, patients are screened for depressive symptoms at various intervals throughout their cancer care using standard and validated screening tools such as the Distress Thermometer [345] or the 2-item Patient Health Questionnaire (PHQ-2) [249]. The goal of this phase is to proactively screen patients for depression at key points in their cancer care and to engage them in appropriate psychosocial care to address potential barriers to engagement in cancer care and to improve health outcomes. When a patient reports depressive symptoms above a critical threshold, the care team refers the patient to the appropriate BHP for a thorough assessment: "*The information from the distress screening device is entered [into the EHR]. If they've hit a certain threshold, it triggers an alert to social work. And then social work calls and follows up with them.*" (BHP3). Patients with pre-existing depression that require coordination with an outside psychosocial care provider or continued management of their condition are also referred to BHPs: "*If a medication seems to be indicated or if they have a more complex mental health background, then psychiatry or psychology would be an appropriate referral. But it really should happen after they're referred to social work.*" (BPH9).

Phase 2: After patients are identified as potentially depressed and referred to a BHP, they move to the **initial psychosocial assessment, diagnosis, and rapport building** phase. The goal of this phase is to gain a deeper understanding of the whole patient, assess the needs of the patient, build trust and a working relationship, conduct a diagnostic assessment, and provide resources and a depression treatment plan. The patient meets with a BHP who performs a full psychosocial assessment of the patient by using validated tools to establish baseline symptom severity (e.g., 9-item Patient Health Questionnaire (PHQ-9) [223], 7-item Generalized Anxiety Disorder (GAD-7) [385]), by eliciting background information (e.g., family and medical history, psychiatric and substance use history, financial background, social support, practical needs), and by conducting a comprehensive clinical evaluation to establish a diagnosis prior to initiating treatment [1]. Through one or more interactions with the patient, the BHP builds rapport with the patient "*to best know them and support them*" (BHP4), "*to open up or unpack what [BHPs] need to track them*" (BHP6), "*for patients to begin to trust more and join with [BHPs]*" (BHP5), and "*to get a patient buy-in [and know] whether or not they're gonna continue to ask for your support*" (MP2). This phase also provides an opportunity for patients to receive psycho-education and supportive resources: "*One of the things I try and do with patients is also let them know the different supportive care resources here, and that any of them, whether it's psychiatry, psychology... it's okay for them to see any of... It's really giving 'em permission.*" (BHP6).

Phase 3: After patients have worked with BHPs and oriented themselves to available psychosocial support, the next phase of the psychosocial care journey involves **active depression treatment**. The goal of this phase is to target identified problem areas from the previous phase and to achieve target patient goals (e.g., remission of depression as indicated by low PHQ-9 scores, and cancer treatment adherence). BHPs provide brief evidence-based treatments (e.g., problem-solving therapy, motivational interviewing, cognitive behavioral therapy (CBT), behavioral activation (BA), acceptance and commitment therapy) during scheduled sessions (e.g., weekly, bi-weekly) based on the needs of the patients: "*I will be doing BA with him mostly because of the way he and his wife talked about his activity level... he didn't identify with the word depressed, but he said, 'I just don't feel like doing anything.' Like the lack of motivation, and he's less talking about like, 'I'm thinking about X, Y and Z,' where... if he's having thoughts that are getting in the way, that would make a difference. I would do... be doing CBT with him.*" (BHP4). Patients regularly complete patient-reported outcomes (e.g., PHQ-9, GAD-7) and perform tasks or assignments between sessions for continuous care and positive reinforcement: "*I really do enjoy opening the sticker... the actual peeling of the sticker, and putting it next to my whatever it, my goal... whatever it was. It makes you feel good. So she actually prescribed that you should get this activity journal, log it, put the sticker next to it.*" (Pt8). Patients that need psychotropic medication management or that have additional co-morbidities (e.g., co-morbid anxiety, substance use disorders) are typically also seen by psychiatrists.

Throughout active depression treatment, BHPs monitor the progress of patients using assessment tools and their clinical judgment: "*I kind of rely on the [assessment] tool to sort of be the indicator of how things are going.*" (BHP6). BHPs systematically monitor all patients under their care, proactively reach out to patients when necessary, and conduct regular caseload reviews with psychiatric consultants to make treatment adjustments. For patients who are discharged after completing their cancer treatments or who move away from the cancer center, BHPs facilitate referrals to and coordination with community providers for continued engagement in care: "*There are some patients who have serious mental illness*[1] *and I actually think need to be in a community mental health center where they have a case manager. And so earlier on I will try to get them involved as soon as possible once they're like physically well enough to be able to travel to that place. And then we'll work together for a while and then I'll try to sort of let the community mental health agency take over.*" (MP2).

Phase 4: When a patient's depressive symptoms substantially improve and the treatment target has been reached (e.g., depression remission), patients move to the **maintenance and relapse prevention planning** phase. The goal of this phase is for patients to achieve long-term remission of depression and maximize self-efficacy. Patients work with their BHPs to establish a relapse prevention plan that includes identifying early warning signs (e.g., changes in sleep, mood), establishing effective coping strategies (e.g., taking medications, reducing stress, engaging in pleasurable activities), and contacting BHPs if symptoms re-emerge [193].

---

[1] This discussion was in reference to a patient with major depression that persisted beyond active cancer care.

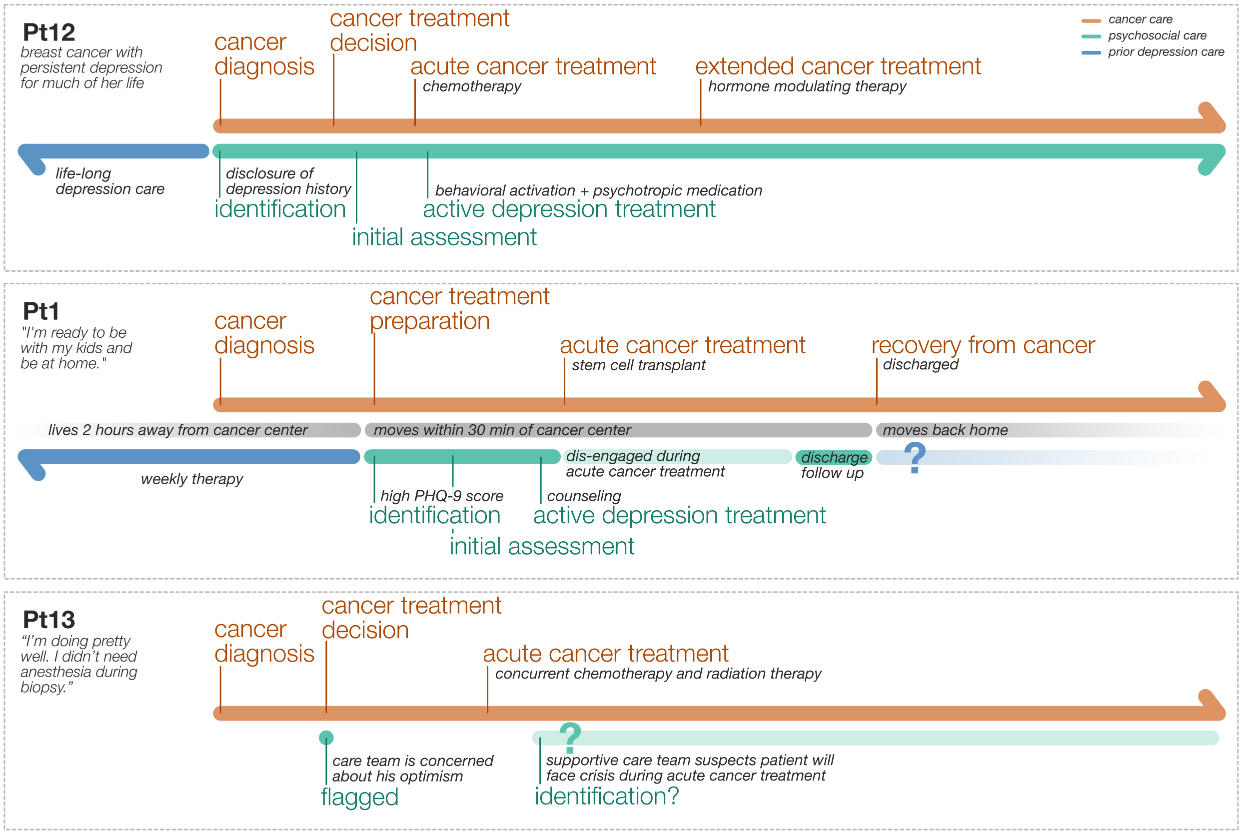


Figure 4.3: Illustration of parallel journeys for three patients with cancer and depression. Different patients engaged in psychosocial care at different phases of their cancer care. These patient journeys were constructed from a medical provider interview (Pt12 described by MP2), a patient interview (Pt1), and observation of a BHP (Pt13 session with BHP2).

As with a cancer care journey, the psychosocial care journey alone is insufficient for characterizing the care experiences of patients with cancer and depression. For example, depression care could continue beyond cancer treatment, but because behavioral health services are often embedded in and limited by cancer center resources, patients who move away after active cancer treatment or are discharged from the cancer center may need to search for new providers. Recurrence of depression could place patients back into the psychosocial care journey, and such patients may be engaged in different phases of cancer care, introducing complications of coordinating multiple care resources.

Parallel Journeys Framework

Both the cancer care journey and the psychosocial care journey were alone limiting in characterizing the challenges of patients with cancer and depression, in part because the combination of the two journeys further complicates the patient experience by exacerbating the characteristic challenges of each of them. To address this limitation, we developed the concept of a *parallel journeys framework* as a conceptual design framework to examine these journeys together and to help organize our data. To illustrate this concept of parallel journeys, we describe the parallel journeys of three patients extracted from our qualitative data. Pt12's experiences were described by MP2 during her interview. Pt1 discussed her own experiences during her interview. Pt13's experiences were constructed based on our observation while shadowing BHP2. We also listened to BHP2 discuss Pt13 with a nutritionist, and we asked clarifying questions about Pt13 to BHP2 outside of the session.

Prolonged engagement in psychiatric care: Pt12 (Figure 4.3 top) has experienced life-long depression prior to her breast cancer diagnosis. Because of the complexity of her depression and her need for psychiatric support, she is immediately referred to see a psychiatrist. Upon completion of her chemotherapy, she is placed on hormone-modulating therapy for the next five years with regular monitoring for the recurrence of cancer. Because she is still being treated by the cancer center, she remains actively engaged in the psychiatric care that the cancer center provides once every two months.

Pt12's hormone-modulating therapy will last five years, during which time she will be eligible to see the psychiatrist at the cancer center. Although the cancer diagnosis had initially caused her depression to be severely decompensated, the psychiatrist recommends she see an outside BHP regularly because the psychiatrist does not have more frequent appointment availability: "*From a therapeutic standpoint... she would probably do better with more frequent therapy. We just simply don't have the availability.*" The length of her cancer treatment puts a strain on her depression as well as on the psychosocial care team at the cancer center which is obligated to treat existing patients while turning away new patients due to an overbooked schedule.

Abrupt termination of depression care: Pt1 (Figure 4.3 middle) moved across the state to receive a stem cell transplant[2]. Prior to her cancer diagnosis, she normally saw a therapist every week to manage her depression. The standard preparation phase of her transplant process includes a full psychosocial assessment by a BHP, so her initial PHQ-9 assessment revealed a high level of depression and a need for continued psychosocial follow-up. Her transplant was urgent, so she was disengaged from her depression care with no support during three weeks of acute cancer treatment. Her follow-up session with the BHP happened when she was ready to be discharged from the hospital. She moved back home as soon as the cancer treatment was completed, and she was not able to develop an ongoing relationship with the BHP at the cancer center or to receive a proper transition plan.

Pt1 faced two challenges during her depression care. First, her depression care was paused during transplant: "*if my circumstances were a little different, because my transplant was a little rushed... I probably would've scheduled something with [BHP] every week.*" Second, her follow-up session with the BHP was focused on her wanting to move back home earlier than recommended, and she left the cancer center without a proper transition plan to continue her depression care. She lives two hours away from the cancer center and wished for a phone follow up: "*she could've just called me for a few minutes just to like say, check in, and see how I was doing.*" Her abrupt transition out of depression care illustrates some of the missed opportunities that contribute to the under-treatment of depression in patients with cancer.

Anticipated crises during active cancer treatment: Pt13 (Figure 4.3 bottom) was diagnosed with head and neck cancer with a part-time income that barely covered his living expenses. He was referred to the BHP by the oncology staff because his family seemed to be more distressed than the patient himself about the diagnosis. He was scheduled to receive concurrent radiation and chemotherapy treatment which is notoriously difficult on the body. However, he had an optimistic view about his treatment and was willing to remain employed throughout treatment to support himself. Although his depression and anxiety assessment scores did not currently indicate

---

[2] Patients undergoing bone marrow or stem cell transplant are asked to live within a 30-minute drive of the cancer center.

clinically significant symptoms, the BHP flagged the patient for monitoring because she suspected he will need help when the intensity of the treatment starts to affect his physical capacity to work and may negatively impact his mental health.[3]

Pt13 had not yet been diagnosed with depression, but his supportive care team suspected that the severity of his cancer and treatment and possible subsequent unemployment would put him at high risk for depression[4]. If he develops depressive symptoms during his acute cancer treatment, as anticipated by his care team, he would need to be identified through a screening process to receive depression care. Without a proper screening process, his depressive symptoms could be easily overlooked because oncology providers are not focused on diagnosing mental health disorders and male patients often do not self-disclose their depressive symptoms [382, 307, 273].

As Figure 4.3 illustrates, the care experiences of these three patients are vastly different from each other. These experiences represent not only three distinct ways that patients with cancer engaged or would engage in their depression care, but also highlight unique challenges brought on by the interaction between cancer care and depression care. Patients could be engaged in depression care for a prolonged period due to the length of their cancer treatment (Pt12), or depression care could abruptly end because cancer treatment is completed (Pt1). Acute cancer treatment could lead to disengagement from depression care (Pt1), or it could cause depressive symptoms (Pt13).

4.1.3 Two Journeys, Additional Challenges

The three patient experiences described above demonstrate that different patients engage in depression care at different phases of their cancer care according to their cancer diagnosis, treatments, physical and emotional conditions, and personal and environmental circumstances before, during, and after their cancer care journey. In order to consider the challenges that patients with cancer and depression navigate when these parallel journeys interact, we examine the cancer care journey described by Jacobs et al. [199, 198] alongside the psychosocial care journey as a

---

[3] Our study did not follow this patient beyond this observation.

[4] Patients with head and neck cancer have particularly high rates of depression and suicide [280, 334, 300].

design framework, which we refer to as a *parallel journeys framework*. This framework is used to organize our findings, specifically the impact on patients and their care team (e.g., BHPs, psychiatrists, oncologists) due to: (1) the interactions and sometimes competition between the two care journeys, and (2) challenges introduced by the current cancer-centric processes of care. Table 4.2 demonstrates the use of this framework by providing examples of challenges that we observed at each intersection point between the two care journeys. Although we did not observe challenges for every possible intersection, these examples show how each stage of the psychosocial care journey produces challenges in at least one stage of the cancer care journey, and vice versa.

Such a framework can also be useful in identifying challenges that may not be directly observed in data. For example, we did not explicitly interview patients who developed depression after their cancer treatment. However, one BHP explained the importance of monitoring patients after cancer treatment [112, 315, 288]. When her previous transplant center enforced conservative post-treatment guidelines (i.e., patients are sequestered at home for 6 months to a year), she would strategize with patients on productive use of time at home to counter feelings of depression that may accompany social isolation (BHP5).

We also found that interactions between the two journeys impact the stakeholders differently, suggesting the third dimension of *stakeholder role* in the organization in Table 4.2. Because our goal is to design technology for providing *behavioral health care* and for supporting *patients and their care team*, we present these multi-dimensional challenges primarily by phases of the psychosocial care journey and according to different stakeholder perspectives. Then, in each subsection, we highlight specific interactions with phases of the cancer care journey.

Psychosocial Care Journey: Identification of Patients with Depression

We observed a number of challenges in detecting depression in patients with cancer (Table 4.2, column 1). The best practice guideline for patient identification is to *longitudinally* screen for patient distress throughout their cancer journey using validated assessment tools [316, 137, 295], necessary because depression can be undetected in patients with cancer who may develop it at any point during their cancer journey and because oncology providers are not focused on identifying

Table 4.2: Challenges at the intersections between psychosocial care journeys and cancer care journeys. Columns represent phases of a psychosocial care journey, and rows represent phases of a cancer care journey. At the intersection of the two journeys, we provide an example challenge identified in our data, noting that additional unobserved challenges may exist.

	Identification of patients with depression	**Initial assessment, diagnosis, and rapport building**	**Active depression treatment**	**Maintenance and relapse prevention planning**
Cancer screening and diagnosis	Information overload from both cancer diagnosis and supportive care resources (Pt10)	Patients under-report depressive symptoms during initial assessment (BHP1)	Depression is decompensated with a cancer diagnosis (MP2)	Cancer diagnosis can exacerbate existing stressors and risk factors (MP1)
Initial information -seeking	Patients hide depression following a cancer diagnosis (MP1)	Patients may get tired of paperwork or refuse initial assessment (MP2)	<unobserved>	<unobserved>
Acute cancer care and treatment	Oncology staff may not screen for depression (A1)	Initial assessment can be difficult when cancer-related stressors take precedence (BHP3)	Lack of energy impedes depression treatment (MP5)	<unobserved>
No evidence of disease	Depression may not be identified when not seeing oncologists or BHPs (BHP5)	<unobserved>	Patients are not transitioned to community providers or lose access to BHPs following cancer treatment (Pt1)	<unobserved>

mental health conditions [382, 307, 273]. With inadequate screening, BHPs are confronted with emotional crises that arise from underlying depression. However, we observed that all three sites did not administer distress screening longitudinally (e.g., PHQ-2). The screening was instead administered at the initial oncology consultation and sometimes at other pivotal moments (e.g., after treatment, at relapse), a timeline defined by the cancer journey rather than the psychosocial care journey.

Patient Perspective Some patients hesitated to report depressive symptoms due to unwanted consequences: "*And if you're honest and you say that you consider depression, all the brakes come on and that's a bad thing. Because that makes people- or it makes me not wanna rec- put down that I'm depressed. It's gonna get negative attention. It doesn't get positive.*" (Pt5). It was common for patients with cancer to describe feeling pressured to be strong. Patients we observed felt they had to hide their vulnerability or even their disease from their children or family and did not have a safe place to be vulnerable. One patient said, "*Oh my gosh. It took me forever to really admit it, 'cause I was ... I didn't have to put on bracelets that said, ... you're strong. And I was just, like, I can't be depressed, I can't be depressed. And it just finally, it was, like, I give up.*" (Pt8). Some patients normalized depression as an obvious side effect of cancer: "*[They] laugh it off and say, 'well I've got cancer of course I'm depressed.'*" (MP4). Some patients expect depression to go away when cancer goes away: "*They say 'Well, this is just related to my cancer. I treat my cancer, and you know... I won't have to worry about my mood.'*" (MP5).

Because cancer demands so many resources from everyone surrounding a patient, patients often felt guilty and were hesitant to seek out help when they needed it: "*I don't have employment anymore, I get charity from the hospital, so they pay for me to see all these people and so there's that hesitance too, where it's like, I'm already aware of how much they're paying. I mean, do I wanna ask for more? There are other patients here that also need resources. How much should go to me and... Balancing that and what that feels like psychologically as well.*" (Pt11).

Behavioral Health Provider Perspective When patients are not adequately screened and identified, BHPs commented that patient distress could "*explod[e] into crisis*" (BHP1) at any moment and that needing to manage to such crises as they unfold is not an ideal situation for bringing patients into depression care. Social workers also described being "*dumping grounds*" (A1) for all problems that the oncologists cannot address: "*It's common if somebody cries, that some of them come running to us. They get uncomfortable with a lot of emotion in the room.*" (BHP8). One administrator said, "*In many places I've been, that was kind of like, if you didn't know what to do with it, send it to the social worker. (laughs). Yeah. That's not okay.*" (A1). Even though the BHPs are trained to assess and treat anxiety and depression within the oncology context, some providers bypass the BHP and refer patients directly to psychiatry: "*We're still struggling I think in some ways in social work overall, to emphasize our role to the team as being the primary resource for mental health care versus psychiatry. We have the skills to assess and treat, you know, anxiety and depression within the oncology context. So, there's no reason to bump up to a higher level.*" (BHP9).

Medical Provider Perspective Primary care visits and oncology visits are common points in cancer care where depression is identified. All three oncologists we interviewed reported that they identify patients with depression through the use of their "*gut feelings most of the time*" (MP1) and from patients describing depressive symptoms (i.e., apathy, fatigue, forgetfulness that cannot be attributed to current treatment or medication) (A1) or due to patients "*break[ing] into tears*" (MP3) during their follow-up visits. However, "*sensing*" distress is difficult for providers because they lack sufficient psychiatric training, some patients are not forthcoming about depression, and distressed patients often present to providers with physical symptoms rather than mental health issues [137]: "*I will be with the patient for months on end and then it's not until I make a referral to palliative [care] that palliative says 'Hey, did you know your patient's depressed?' I'm like 'Oh, how did I miss that?'* " (MP1). One provider said, "*My suspicion would be that patients may minimize their symptoms a little bit to me. I think there's a little bit of that they don't want me to feel disappointed that they're not responding [to medication].*" (MP5).

We also observed that oncology nursing staff may be more attentive to patient depressive symptoms than medical oncologists: "*When I first started working here, a nurse brought to my attention a patient who they thought could benefit and when the oncologist was approached about putting in a referral for this patient, he remarked, 'Well, why? She's one of my happiest patients.' And, in actuality, this patient was incredibly depressed.*" (BHP3). However, the staff was under pressure to improve the efficiency of oncology care processes, which limited the opportunity to take time to administer depression screening or adequately triage patient psychosocial needs for proper referral: "*There's a lot of pressure from providers to get in and out of the room so that they can get into the room and see patients. And so they are there always pressuring the MAs, LPNs, you know, to* *do things quickly and get out.*" (A1).

Psychosocial Care Journey: Initial Psychosocial Assessment, Diagnosis, and Rapport Building

We also identified challenges in the assessment phase of the psychosocial care journey, either associated with active cancer treatment (Table 4.2, column 2) or based on a variety of needs that patients had at different phases of their cancer care. Because cancer places such high demand on financial, physical, and practical resources [229], patients were more likely to drop out or not show up to the first appointment with their BHP or to occupy most of the initial assessment appointment with discussion around their practical needs (e.g., financial aid, transportation, housing) instead of treating their depression.

Patient Perspective The first visit is an opportunity for a patient to express their emotional support needs and to negotiate what kind of care they can expect from the follow-up sessions: "*[we discussed] things that I want to address or that I'm struggling with ... or what I'm wanting out of the sessions.*" (Pt11). To some patients, it was the first time they felt validated: "*She came to me when I was first admitted to the hospital. And, she's been a godsend. Just because she would... there's some validation that you're going through something.*" (Pt5). We also observed that the resources and support that patients received during this phase depended on where the patient is in their cancer care journey. For patients that are just entering cancer care or are in an unfamiliar environment far away from home, the initial assessment helped patients be connected with someone: "*When I first went to [treatment city], I was by myself for a really long time, so I didn't really have a whole lot of people to talk to. So, getting to sit down and talk to [BHP] about how things were going and how I was feeling about my whole transplant was really helpful for me.*" (Pt1). For patients who have been fighting cancer for an extended period of time, the initial psychosocial session also helped patients establish a safe space to be vulnerable when the usual cancer culture encourages them to "*be strong and keep fighting*" (BHP1). For patients with advanced cancer, it provided an opportunity to have a conversation about seemingly awkward but important topics such as the power of attorney, end-of-life directives, or options for physician-assisted hastened death.

Patients were sometimes not aware they were referred to see a BHP: "*Sometimes they (patients) didn't know that they were scheduled for us (BHPs) so they're like, 'What the heck is this about?'* " (BHP1). Patients may cancel or skip initial appointments because there has not been an established relationship: "*Our no shows are more common with patients that have never met us in the first place, I think partially because we don't have a relationship.*" (MP2).

Behavioral Health Provider Perspective When BHPs are introduced to a patient during a crisis moment, BHPs often perceive the need to drop everything to attend to those crises. In addition, patients are often overwhelmed with the cancer-related issue at hand such that BHPs cannot perform a full psychosocial assessment, share psycho-education or resources in a digestible manner, or identify goals or a treatment plan for future sessions. One BHP commented "*because you're dealing with the crisis and the emotional fallout, ... I think that sometimes our assessments are jumbled and not thorough. And you can kind of try and backtrack at the next meeting to do a better assessment, but that doesn't always happen either.*" (BHP1). Another said "*when you're going into a crisis moment, it's hard to just say, 'Oh, wait. Let me give you this survey.' (laughs)*" (BHP2).

Cancer places such a high financial burden on patients that BHPs spend much of their time addressing patients' practical needs instead of their underlying depression: "*They're always having to deal with financial issues, or this person needs housing, or lodging, or transportation.*" (MP5). Social workers often are not able to establish themselves as BHPs during the initial assessment: "*I think especially medical social workers, just get dragged into the practical stuff and never use our other [behavioral therapy] skills. And so I just feel like my skills are really rusty.*" (BHP1).

Psychosocial Care Journey: Active Depression Treatment

The intersection between active depression treatment and various stages of cancer care presents additional challenges (Table 4.2, column 3). During active depression treatment, BHPs must be flexible to frequently-shifting patient circumstances associated with cancer and oncology treatment. Depressive episodes vary in both length [384] and rate of recurrence [71], and cancer diagnoses and treatments present unpredictable logistical [82] and health [171] challenges. This leads to significant challenges for patients, BHPs, and medical providers who need to navigate the dual journeys of depression and cancer care. We found that these challenges can arise during the acute treatment and recovery phases of cancer care.

Patient Perspective Many patients with cancer and depression may disengage from depression treatment due to fatigue, which is both a symptom of depression and a common side effect of cancer treatment. The heavy burden of oncology and psychosocial appointments may be especially overwhelming for depressed patients who lack motivation or hope, reducing patient willingness to engage in one treatment or the other: "*I guess I'm just tired of all these appointments happening, I'm like can I get three weeks free without an appointment, you know?*" (Pt3). Patients may also feel physically or cognitively drained when going through chemotherapy, losing motivation to seek out depression treatment: "*It's hard to prioritize seeking that out if you're already dealing with something else, because it's like, 'Man, I got treatment...I'm dealing with the side effects of the last treatment.'*" (Pt4). In some cases, external resources such as books, apps, and support groups can be helpful to patients in this phase: "*I joined this—on Instagram—a group of metastatic breast cancer people, and I look at all this thing. All this, thousands of women around the world, just sharing their thing. And, it gives me hope...*" (Pt10).

Another challenge for patients arose out of a general lack of documentation of psychosocial session outcomes: "*Other providers have discharge notes through [registry system]. I don't have that. Sending home discharge notes is impractical because I can't type during the session.*" (BHP3). Often, patients may have memory challenges while going through chemotherapy: "*I did have chemo and sometimes that brain, you know, remembering dates, exact times, a little hard. And side effects from a lot of the stuff I take and I hate that my memory, that's one thing that I noticed had changed. My memory.*" (Pt3). Patients therefore may have trouble keeping track of session outcomes and assignments, and some suggested that tools to help manage that information could be helpful: "*It would be, it could be a useful tool. I've never really thought about it. He [psychologist] kind of references back like, how I was feeling at the past visit. You know, a little bit but and maybe he keeps* *track of it but I don't really even remember what happened yesterday. I'd have to look back up in the calendar to see when it was...Yeah, that's pretty cool. Some additional tool that's specific to mental health? That tracks your fears, how you've been feeling? I don't know, that'd be cool.*" (Pt2). Such tracking may empower patients to develop deeper insights about their own behavior: "*Usually it's the patients where their mood is all over the place. Or they really haven't yet linked the fact that activities or physical sensations link to their mood, and I'm trying to encourage that link—the association in their head.*" (BHP1).

Behavioral Health Provider Perspective During the acute cancer treatment phase, patients may encounter unpredictable crises that derail depression treatment: "*There will be times where we have one session, and we've done a lot of Behavioral Activation stuff, but then the next time they come in, maybe they've gotten some really bad news or something and they're just really distraught, and that session, rather than following-up with the Behavioral Activation, we're kind of doing more of a kind of crisis management.*" (BHP3). Patients may choose to end their depression treatment prematurely when their oncology treatment is finished, despite patients still being symptomatic from depression or despite BHP's belief they could benefit from further sessions or are at high risk of recurrence: "*We are, in my mind, not specifically mental health like you go and you're coming because you want to get mental health treatment. We are coming to support you to get you through the treatment, and so if you're starting to feel better because physically you're feeling better...I can't really justify telling them like, 'Oh, you must meet me again.'*" (BHP4). Clinic social workers may have limited capacity to see patients discharged after completing active cancer treatment, leaving patients to manage their own depression care: "*So I don't really know how anyone can help me at this point. I'm still just trying to figure out how I can help myself.*" (Pt1).

Another challenge emerges due to the need to coordinate between oncology and psychosocial appointments. Appointments with BHPs are frequently scheduled on the same day as cancer treatments, because of the convenience for a patient. However, patient psychosocial needs may not easily fit into this schedule, and it may be difficult for a patient to come in for appointments outside scheduled oncology treatments: "*So, to ask them to come for an exceptional visit is very* *hard for them sometimes... I try to schedule it when it's convenient for the patient because... they've just finished some treatment and they don't have another stretch of treatment for two weeks, chances of them come back in here when they're not feeling well is slim to none.*" (BHP5).

Medical Provider Perspective Just as frequent oncology appointments may make adherence to depression treatment more difficult, depressive symptoms can likewise impair oncology treatment: "*Usually I just let the patients know I am worried about this and make it clear that this could interfere with the main treatment, which is the oncology-focused treatment that I'm trying to achieve for them.*" (MP1). In particular, depression could potentially lead to poor decision-making around oncology care: "*I get concerned at times that there are patients who are making bad decisions about that care. Possibly because of depression...I mean, there are times frankly when I've had patients who had a decision-making process so flawed, and they're making such a bad decision that could have really negative impact on their outcome.*" (MP3).

Nearly all participants, including both oncologists and patients, reported the lack of available appointment slots among the behavioral and psychiatric care providers as one of the main barriers to accessing mental health care: "*The counseling is something that I would love to have for every patient. But the problem becomes, how do I arrange that? [psychologist] is one person. He can't see absolutely everybody, and then when he does, and he wants to do ongoing CBT. It's very hard to schedule appointments with people.*" (MP5). When access to psychiatry is limited, the BHPs, primary care physicians, and oncologists often needed to collaborate to determine the appropriate set of psychotropic medications. This was not an ideal solution, given that none of these providers were properly trained to manage such situations: "*I'm a medical oncologist so I feel a little bit over my head (laughs) in that situation. 'Cause I wasn't trained to manage bipolar disorder.*" (MP2).

Psychosocial Care Journey: Maintenance and Relapse Prevention Plan

We were not able to observe any patients moving into the maintenance phase, possibly because patients were discharged from the cancer center before their target depression treatment goals were reached. Instead, patients simply disengaged from depression care because they were discharged from cancer centers (Table 4.2, column 4). There were also some efforts to connect patients with community mental health specialists to continue their depression treatment outside of the cancer centers.

Patient Perspective Some patients needed to manage their own depression, due to a lack of proper hand-off to appropriate community BHPs that patients can access outside of the cancer centers: "*I'm trying to figure out what I'm supposed to do with myself and what kind of person I'm supposed to be in this next year. This isn't really anything that the doctors really told me would happen, that I would feel so lost and unsure what to do next.*" (Pt1). In some cases, patients felt reluctant to seek treatment elsewhere: "*She [patient] also feels a sense of safety at [urban site], and so going to a new place, it's scary for her.*" (MP2). This presents a challenge, as cancer center behavioral health resources may not have the capacity to continue treatment after a patient's cancer treatment is finished: "*I had a patient once who was angry, but she hadn't had cancer in 12 years when she came here and she wanted to see a social worker. And I said, that's just not, you know, it's not possible.*" (BHP5). One patient explicitly rejected receiving continued depression treatment at the cancer center: "*I just associate this place with (laughs) not good feelings, I don't know. It sounds mean, but I mean it's just... the least amount of time you have to be up here, the better.*" (Pt8).

Many patients, especially those needing special treatment that their local or rural cancer center did not offer (e.g., stem cell transplant), had to relocate or travel for hours for access to the urban cancer center. Unlike psychosocial care in primary care settings where patients are likely to remain in the same primary care facility in their local communities, some patients would need to travel for hours to continue their depression treatment: "*It's challenging when they don't live here... It's challenging that they can't get here.*" (BHP5).

Behavioral Health Provider Perspective As much as BHPs and psychiatrists wanted to continue treating patients, they also reported their skills and expertise should be used for those who require psycho-oncology support: "*This is one of our struggles. Because she wants to see me. I think she gets something out of our visits. I could probably see her for years (laughs) and she has* *breast cancer, and the breast oncologists like us to take care of their patients long term. But she doesn't necessarily have to have a provider who specializes in psycho-oncology at this point unlike some of our patients. So I think each provider has a different opinion, but for me, I hope to be able to graduate her because I want to open up my clinic to new patients.*" (MP2).

Although patients were welcome to reach out to their BHPs during active cancer treatment, BHPs were also limited to only seeing patients in active cancer care: "*We have a limited number of appointments that we allow our behavioral health providers to see after treatment is done, because we don't have the staff here to keep those patients coming back day after day.*" (A2). One site partnered with a startup that offers to match patients to BHPs, but there were exclusions: "*We sometimes have to find a resource for patients if they're outside of that area or [on] Medicaid.*" (A3).

4.1.4 Technology Design Opportunities for Collaborative Care for Depression and Cancer

We have presented specific challenges that patients with cancer and depression encounter. Informed by these challenges, we have also identified design opportunities aimed at improving collaborative care for depression in cancer settings, specifically opportunities that can be addressed through the use of technology. Although some of the identified challenges may not be unique to this setting, collecting insights with this population allowed us to more saliently observe the particular issues facing those undergoing both cancer care journeys and psychosocial care journeys, in turn allowing us to better prioritize functionalities and design opportunities. Whereas prior sections are organized by phases of the psychosocial care journey, this section presents a set of technology design opportunities and how each opportunity can improve the care experiences across multiple phases of the cancer care journey and the psychosocial care journey.

Provide tools for and around self-assessment

From a patient perspective, we observed that patients in active cancer treatment may experience physical side-effects of the treatment (e.g., chemotherapy) and difficulties with transportation, both of which may make it more difficult to attend appointments. Furthermore, some patients in active treatment noted that they had significant downtime in the infusion suite with little to do, suggesting another opportunity for self-assessment activities. Finally, patients nearing the end of treatment described being tired of making trips to their treatment sites, suggesting they may also benefit from home self-assessment. Providing patients with self-assessment tools could help patients stay engaged in care through remote monitoring when in-person appointments are less desirable or feasible.

For example, systems could allow patients to perform technology-supported self-assessments (e.g., universal distress screening or depression and anxiety assessment performed through electronic surveys, apps, or IVR technologies). The resulting data could be directly integrated with an EHR system, further reducing the time spent on manually entering such data. Self-assessments could also be configured to be administered at an appropriate frequency, thus avoiding a need for BHPs to determine whether the assessment needs to be administered before every patient visit. Such a system would allow the frequency of assessments to be decoupled from the appointment schedule to accommodate patients at various phases of their cancer care journey.

From a related provider perspective, we also observed that a patient's cancer journey creates additional challenges for BHPs responsible for psychosocial care. BHPs play a critical role in coordinating patient care across many providers and supporting patient needs from multiple perspectives. Treatment of patient depression during the limited time available within sessions, therefore, competes with other tasks that are directly (e.g., responding to patient voicemails, handling crises) or indirectly (e.g., contacting Medicaid case managers, reviewing financial aid applications) related to patient depression. Many of these tasks are typically organized through memory, hand-written notes, or email/calendar reminders. We also saw that patient depression care is often neglected when BHP resources are constrained [286]. Technology improvements that help BHPs manage competing demands, or even improvements that allow more effective use of limited time, could allow BHPs to focus on activities critical for patient management and for providing psychosocial care.

One such opportunity is to give BHPs access to patient self-assessment data, thereby reducing logistical burdens and allowing BHPs to focus on using session time to better deliver depression treatment itself. Systems can also reduce inefficiencies in managing and distributing patient care resources (e.g., sifting through binders of care materials, photocopying resources, walking back to their desks to find materials) through a repository of resource materials that all BHPs have access to and which can be emailed or printed for patients as needed during a session. Similarly, systems can also allow BHPs to configure, schedule, and organize action items with priorities, reminders, and deadlines. Such electronic assessment systems have been shown to be feasible and effective in cancer settings [138, 136, 44, 43]. In addition to being consistent with recommendations of prior research, our results characterize these challenges as resulting from the parallel journeys of patients with co-morbid cancer and depression. Our examination of the perspectives of multiple stakeholders shows this is a key priority for both patients and providers in this setting.

Provide tools for population-based patient monitoring

We found challenges in a patient's cancer journey can often influence patients to disengage from their depression care. Monitoring patient progress is a critical component to ensure an appropriate level of patient engagement in their depression care. With appropriate monitoring in place, providers can both: (1) detect when patients are not improving as expected and therefore make appropriate adjustments to treatment, and (2) identify patients who are disengaged from care and therefore initiate proactive outreach. However, all three sites lacked appropriate and efficient tooling support for population-based analysis and monitoring of patients. One site leveraged an Excel-based patient registry to systematically follow all patients, but the corresponding workflow created duplication of efforts (e.g., the registry was not directly integrated into the EHR), which discouraged proactive monitoring of patients. To streamline population-based patient monitoring, systems could allow patient self-assessment data to be directly entered into the EHR (i.e., reducing the burden of data entry) and could also provide a summary dashboard for easy access to pertinent information about a provider's population of patients. Although prior research has proposed EHR systems that collect data from clinical registries [54] or aggregate information from various databases [366], this solution would build upon our suggestion for patient self-assessment to further address the needs of this patient population. For example, such a system could help track patients undergoing active cancer treatment, who may “fall through the cracks” when cancer care is prioritized over depression treatment. Furthermore, given that patients who have completed their cancer treatment may lose access to the cancer care center, an electronic registry could trigger action items or send reminders to BHPs to follow up with patients who have been discharged or moved away but still require depression care (e.g., to help ensure appropriate transition plans).

Patients we interviewed reported frequently checking their progress toward cancer recovery by logging into their EHR portal and viewing historical lab results and provider notes. The same system could also help patients reflect on their progress with depression treatment, which could help raise awareness of the importance of treating their depression. A system might provide rewards and achievements intended to increase self-efficacy or might guide patients with appropriate interventions and psycho-education materials.

Provide access to digitally translated evidence-based psychosocial interventions

We observed that patients often have significant periods of downtime across multiple stages of the cancer care journey. For example, during active treatment, patients may spend several hours at a time in the infusion suite. Furthermore, stem cell transplant patients going through recovery may not be allowed travel, left with long periods of time with potentially little to do. For those patients who are also undergoing treatment for depression, these may represent opportunities to provide electronic resources (including tools such as worksheets) for evidence-based psychosocial interventions. By translating behavioral health resources into an engaging digital format that patients can access via their phones, tablets, or computers (e.g., while at home or in an infusion suite), patients could be better supported in their psychosocial care journeys during these long stretches mandated by their cancer care journeys. Easy access to such resources could be especially important for those undergoing active cancer treatment, where we observed that unpredictable crises can be common.

Digital resources for evidence-based psychotherapy have been of interest to researchers for some time [220, 90, 195], including resources developed for anxiety and depression [331]. A tool could aggregate such existing resources and incorporate them into patient treatment (e.g., patient-facing mobile apps could include search features to identify relevant resources), which could further support patients who have difficulty accessing frequent in-person sessions (e.g., for health reasons, for financial reasons, due to a lack of cancer center depression resources). Our parallel journeys framework also suggests that existing resources or new interventions might be organized according to where patients are in their multiple journeys (e.g., presenting additional resources at times in a patient's cancer journey when they are most at risk for depression).

Document shared understanding between patients and providers

Our findings revealed that forgetfulness and fatigue negatively impact patient engagement in care, a challenge that is magnified by the additional difficulties of navigating parallel journeys. Although forgetfulness as a barrier to treatment adherence may seem obvious, we cannot dismiss the fact that the cancer treatment impacts the brain chemistry in a manner that severely reduces the patient's ability to remember (i.e., so-called "chemo brain") [185, 388]. In addition, patients with cancer receive a cocktail of medications with significant toxicities and endure an overwhelming number of hospital appointments that intensify their fatigue. Although depression care sessions often generated action items or assignments for patients, these decisions were not documented or were documented in a manner such that they were not available when the patients needed them (e.g., sticky notes, paper). Furthermore, BHPs sometimes forgot they assigned anything to a patient, and the resulting lack of follow-up accountability and interest from the BHP can lead to a loss of trust and consequent nonadherence by patients. Given that prior research on patient-provider collaborations using personal informatics data has indicated that clear goal-setting is necessary for managing expectations around data collection and reducing information overload for both parties [89], there is an opportunity for a system to help patients and providers manage a shared understanding of depression treatment, especially given the high demands of this population.

Documentation of action items could be facilitated through provider-facing systems that allow quick charting during patient sessions. Such a tool could be designed for transparent collaboration in the documentation of action items between patients and providers during a session, a strategy that can ensure shared understanding and has been found to build trust and reduce miscommunication in patient-provider collaboration around personal informatics data [361, 89]. Patient-facing systems can make action items accessible in a variety of modalities, especially on mobile devices, a possibility supported by one BHP participant: "*people always know where their phone's at. (laughs)*" (BHP3). Such a system would need to be designed in a manner that minimizes cognitive burden, both during a session (i.e., ensuring it enhances collaborative goal-setting rather than distracting from an in-person session) and afterward (e.g., action items could be lost in typical text-heavy after-visit summary).

Such support for patient-provider collaboration around a shared understanding of patient action items could also facilitate fidelity to active components of evidence-based treatments. For example, systems could include activity and mood-tracking components of behavioral activation, reminders for planned pleasant activities, or positive feedback provided either automatically, as scheduled by BHPs, or synchronously by BHPs in response to patient-entered data. Systems can also include brief behavioral strategies (e.g., goal setting, activity scheduling, problem-solving to address barriers) as well as helpful tools (e.g., basic cognitive or behavioral exercises when patients report depressive symptoms). Patient documentation of their completed action items could support a sense of accomplishment and progress, and providers could use data to follow up with patients to reward and encourage positive behavioral changes or to troubleshoot barriers when a patient is not engaging with action items.

Support timely and appropriate communication

Patients reported they experience intense depressive symptoms between sessions or outside of normal clinic hours and feel a need for guidance from their psychosocial providers. These needs may be exacerbated by the unpredictable course of cancer treatment, which may lead to crisis moments throughout the care journey. Patients overwhelmingly wanted a way to reach their providers outside of their sessions. However, patients hesitated to reach out during these crisis moments due to their own guilt of burdening the busy providers, fear of being on hold, or anticipation of increased anxiety from waiting for callbacks. On the other hand, BHPs were hesitant to receive a flood of incoming communication requests and data that they cannot handle:

"*[handling suicide ideation] would be really hard. I mean, there would be, there's, there's just no way to ensure that there could be a timely response to that. And I think, I think that you've probably pointed out the biggest downside of having any kind of, um, electronic communication that is kinda counseling focused.*" (BPH9). Importantly, patients with depression will experience suicidal thoughts whether or not these thoughts have been communicated to care providers, and arguably patient risk is greatest if providers are unaware of patient suicidality. With careful attention to the design of both clinical processes and technologies, a system that facilitates such communication and appropriate clinical intervention may help reduce patient risk for self-harm.

We observed that planning around communication typically happened during the initial assessment, verbally and through the exchange of business cards. Systems can enhance this workflow by documenting the agreed communication plan and making care team contact information readily available in the desired modalities. Prior research has found that online patient portals can significantly improve patient satisfaction with their communication with their providers, based on increased convenience and direct physician response [241]. Building on prior research regarding the advantages and disadvantages of different communication methods [462], our results emphasize a need for shared understanding between patients and providers regarding the appropriate use of and expectations for communication channels. Our parallel journeys framework also provides a lens for understanding how shared understanding regarding communication can break down (e.g., the reluctance of patients near the end of their cancer treatment to reach out for depression support, due to guilt regarding the already significant support received during their cancer journey, thus contributing to the risk of relapse or unsuccessful transition in their psychosocial care journey).

Asynchronous and virtual communication methods could lower the barriers to patients reaching out and could reduce the time commitment required from BHPs (e.g., messaging can be less time-consuming than scheduling and conducting a phone session). A system could allow appropriate expectation-setting by supporting a BHP in configuring the system to respond immediately with a short message (e.g., regarding when a patient might expect a response). While waiting for a BHP's response, patients could interact with other aspects of a system to find information and interventions appropriate for their circumstances. Because patients require a clear understanding of the level of support for crisis management [31], the design and implementation of such a system will require transparency around the boundaries and limits of using the digital system. BHPs can develop safety plans for their patients and educate patients on real-time intervention tools for suicidal ideation which can be accessed through patient-facing systems (e.g., calling 911, a suicide hotline, a crisis text line). By providing patients with in-the-moment access to crisis resources, patients may also be less likely to show up to BHP appointments in crisis and may therefore be better able to focus on their depression treatment.

Improve access to resources

Patients with cancer can be overwhelmed with information, especially during diagnosis and treatment decision phases [216]. Patients in the initial information-seeking phase may have questions around navigational or practical topics (e.g., housing, childcare, financial assistance, volunteer services, scheduling logistics, the course of treatment); as they transition into active treatment, they may have additional questions about other topics (e.g., transportation options). Patients can also be overwhelmed by the emotions they experience throughout their care. We found that patients were given resources and information, but they were not able to internalize them and wanted frequent reminders: "*Even if you didn't get anything in-between and something came up, having that periodic reminder would... You would know that that resource is available if you need it. Versus just a one-time when you first come in.*" (Pt11). To make patient resources more accessible, systems could intelligently recommend appropriate information based on a patient's specific phase of a cancer care journey and of a psychosocial care journey, similar to personalized systems proposed for patients at risk for cardiovascular disease [84]. BHPs can also manage, curate, and configure appropriate sets of resources based on patient needs and make them available in patient-facing systems. Systems can also allow patients to discover relevant content (e.g., psycho-education resources, self-guided interventions) and could include peer recommendations of useful resources from patients in similar circumstances.

4.1.5 Discussion and Future Work

Characterizing the Whole Patient through a Parallel Journeys Framework

The field of psychosocial oncology has made significant progress in acknowledging the association between psychological and physiological challenges encountered by patients with cancer, highlighting the growing awareness and importance of caring for the *whole patient* and the need to target quality of life as the desired goal for patients with cancer [200]. Despite the increased attention to care for the whole patient and the Institute of Medicine's recommendation to integrate psychosocial services into cancer care [302], there still remain 73% of patients with cancer and depression who receive no or inadequate depression treatment [430]. Our findings confirmed known and unmet challenges in providing psychosocial care in cancer settings [135, 286]: (1) Patients normalized depression as a side effect of cancer. (2) Oncologists did not feel they have sufficient training to manage the psychiatric and psychological aspects of cancer. (3) BHPs were overburdened with navigational and practical responsibilities that interfered with their provision of depression care. (4) Organizations lacked the psychiatric and psychosocial resources and processes to adequately support patients' psychosocial needs. In examining our findings, we discovered that the commonality that underlies these unmet challenges is still rooted in treatment settings that prioritize the treatment of the disease process (i.e., cancer) over the holistic needs of the patient as a person.

We have demonstrated that patients with cancer and depression go through their cancer care journey in parallel with their psychosocial care journey. The concept of the parallel journeys framework came from our finding that cancer-centric views, both in terms of clinical staging of cancer or the phases of a cancer care journey, inadequately describe the experiences of patients with cancer and depression whose psychosocial needs do not necessarily align with cancer-centric views. Despite all of the sites we examined having integrated psychosocial care processes, we still observed across all sites an attitude towards depression being a secondary disease, manifested in how the center structured their budget to hire psychiatric staff, how patients prioritized their limited capacity to cancer-related appointments, how oncology staff supported the distress screening process, and how BHPs over-accommodated to fit the cancer care. As our findings revealed, cancer settings still require *raising the priority of mental health care*. Presenting the psychosocial care journey taking place in parallel with the cancer care journey is one step toward better characterizing the whole patient.

In characterizing the whole patient, our work brought cancer care and psychosocial care journeys together to describe the complexity of co-morbidities, representing the needs of a particular "extreme user group" [85] who have the potential to provide deep insights about these issues [412]. We leveraged existing models of care journeys, and we did not evaluate the feasibility of our framework in other co-morbid contexts or using other models of care journeys. However, our parallel journeys framework may generalize to other conditions because it does not dictate either the types of co-morbid conditions or the specific models of the care journey. Understanding exactly how this framework can be used to describe care challenges in other co-morbidities or using other models of care journey remains an opportunity for future work.

Technology to Prevent Patients from Falling through the Cracks

Our study revealed several missed opportunities that left patients with cancer and depression to fall through the cracks and receive less than ideal or no depression care. Through the lens of the parallel journey, our work exposed many of these "cracks" at the intersection of the cancer care journey and the psychosocial care journey. At a high level, the parallel journeys framework brings attention to technology design opportunities to ensure patients with cancer and depression receive adequate depression care.

Each of the four phases of the psychosocial care journey is delineated by a set of concrete and measurable goals, similar to how the cancer care journey is conceptualized (i.e., with the completion of tasks or changes to the physical body demarcating the phases and paths that patients with cancer experience). Our parallel journeys framework highlights the need for tools to *support achieving psychosocial goals* in each phase of the psychosocial care journey to help patients progress toward improved mental health. Given the goals of each phase of psychosocial care, tools also need to be flexible enough to accommodate the needs and challenges of each phase of cancer care. Having concrete goals promotes measurement-based practices and allows the evaluation of potential tools based on the forward progress of each patient.

The parallel journeys framework also directs attention to the need for tools to *support continuity of care* within and across psychosocial care phases and beyond cancer treatment. Burdens of cancer play a significant role in patients disengaging from depression care, forcing BHPs to be flexible and balance the need to address the crises presented at sessions with interventions to address underlying depression. In addition, because psychosocial care service is integrated into the cancer center, patients discharged from cancer care often discontinue psychosocial care. Improved tools that make psychosocial care accessible outside of cancer centers could support patients that would normally disengage from depression care and could support patients that need in-the-moment help in receiving timely communication and interventions.

Ultimately, it is important to consider the broader context when designing specific policy interventions [45]. Although challenges such as high dropout rates or low adherence to depression treatment are not unique to cancer care settings [303], we observed that the difficulties of depression care are compounded by the difficulties of cancer care, including side effects of cancer and its treatment, financial burdens, and logistical burdens that accumulate from cancer treatment. These burdens are felt both by the patients and by the BHPs who are also overwhelmed by the demands placed on them in the cancer setting. There are many existing technology solutions for monitoring mood and tracking activities [89, 85], but the development of effective interventions and technologies in this context must consider the interactions between the various stakeholders and the specific challenges faced by patients on these parallel journeys and their providers.

Limitations

We only recruited patients with depression whose recent PHQ-9 score was greater than 10. These patients were also reachable by their BHPs and the researchers, potentially introducing selection bias towards patients who are more engaged in depression care. Recruiting actively distressed patients also meant we did not interview any patient in the maintenance phase of a psychosocial care journey. Our sample of patients was relatively small, such that we did not have a representative sample for age, race/ethnicity, gender, cancer type, or cancer stage. Our work also excluded caregivers who are often the driving force behind enabling care for patients. Caregivers, as well as other family members, often need psychosocial care themselves due to the high demand and emotional stress that cancer places on relationships surrounding the patient [297]. More research is, therefore, necessary to understand caregiver challenges in supporting patients, their own psychosocial experiences, and their participation in designing patient-facing tools.

Our contextual inquiry method involved observing sessions that were highly sensitive, private, and emotionally charged. Despite the fact that patients were reminded to ignore the observer in the room, it is possible that they behaved differently during the observed sessions. In our journey mapping exercise, we pieced together a single patient's journey from a single stakeholder's description or from our observation notes, either of which could potentially include biased perspectives. We leveraged existing journey paradigms and guidelines for cancer care and psychosocial care to examine patients with cancer and depression and to identify the importance of considering a parallel journeys framework for characterizing their challenges. The phases of the psychosocial care journey are consistent with clinical practice guidelines [19] and therefore have clinical validity for depression care. However, we have not applied the framework to other co-morbidities and there may be limitations on the generalizability of the phases of the psychosocial care journey in other co-morbid contexts.

4.2 DESIGN OF TECHNOLOGY-ENHANCED COLLABORATIVE CARE

Our prior work revealed that the challenges in depression care of patients with co-morbid cancer and depression arise when the needs for depression care conflict with the needs of cancer care. In particular, we found several challenges specific to behavioral activation as a treatment in the context of integrated psycho-oncology care settings. We saw that the burden of cancer (patients' limited physical, cognitive, and emotional capacity due to the illness and associated volatile treatment and stressors) results in patients' limited capacity for participating in therapy and interferes with the treatment plan. Formal behavioral activation was not being achieved during the course of sessions, or only a few highly simplified or adapted components of behavioral activation were being provided as needed. When the action plans are not documented or are not accessible, then the patients are not aware of what they need to do for their treatment or they often forget to adhere to the plan. If the progress towards the plan is not documented, then the BHPs are not aware of what the patient is asked to do over time and the burden of maintaining and bringing back the session and assignment outputs is placed on the patient. Patients were not able to maintain a steady course of treatment when BHPs are overbooked and unable to schedule regular, re-occurring appointments (due to limited personnel and their need to respond to crises) or when patients have unplanned stressors or barriers to meeting regularly with their BHPs. Because BHPs also have other supportive duties and knowledge (e.g., financial or transportation support) and patients have complex needs and demands, the sessions become fragmented and time is spent on addressing these other needs.

We saw that the core aspects of behavioral activation (e.g., mood logging, activity tracking) become burdensome on top of already overwhelmed schedules and cancer treatments, and patients are generally less motivated to adhere to treatments. Successful behavioral activation requires follow-up and review at face-to-face sessions but is not designed to support asynchronous interactions between patients and BHPs between sessions in case patients cannot have regular visits. If the behavioral activation treatment is disrupted, then the BHPs and patients do not have the tools to continue the care through and beyond the transition. Overall, patients with cancer and depression struggle to prioritize depression care in the midst of an overwhelming amount of burden put on them by both diseases, and the care teams struggle to maintain consistent and high-fidelity depression care due to frequent changes required by the progression of both diseases.

From these findings, we identified six technology design opportunities situated in the collaborative care context that could support patients and BHPs to work together to monitor and improve their engagement and depression outcomes. Our research goal is to leverage these opportunities and enhance existing care strategies in improving stakeholder collaboration and engagement in depression care. We first digitize core components and tasks of collaborative care and behavioral activation and present a model for a technology-enhanced collaborative behavioral activation loop (Section 4.2.1). This model is used to inform the design and development of a deployable collaborative behavioral activation system, named SCOPE (Supporting Collaborative Care to Optimize Psychosocial Engagement, Section 4.2.2). We conduct multiple rounds of iterative design interview studies with patients and BHP stakeholders to confirm or make adjustments to our design choices, and the system is deployed to a pilot population of 6 patients, 3 social workers, and 1 psychiatrist (Section 4.2.3). We analyze the interview data from these four rounds of studies to evaluate how our SCOPE system supports the collaborative behavioral activation loop (Section 4.2.4). We then present technical and non-technical opportunities for improvement and future research directions in deploying a technology-enhanced collaborative care system in a real-world, clinical context (Section 4.2.5).

4.2.1 Digitization of Collaborative Care and Behavioral Activation

The technology design opportunities identified from our prior study are the following:

DO1. Provide tools for and around self-assessment
DO2. Provide tools for population-based patient monitoring
DO3. Provide access to digital psychosocial interventions and treatment modalities
DO4. Document shared understanding between patients and providers
DO5. Support timely and appropriate communication
DO6. Improve access to online and community resources

We leverage these opportunities as our guiding principles in envisioning how technology would enhance existing care strategies. Specifically, our prior study probed multiple stakeholders' experiences with integrated psycho-oncology care on two evidence-based strategies proven to be effective in cancer settings: (1) the Collaborative Care Management model [238], and (2) Behavioral Activation [119, 130]. In this section, we explore these two strategies to highlight core components and actions that we could target for the digitization that address the design opportunities identified earlier. We then propose a collaborative behavioral activation loop model to illustrate how technology can support stakeholder engagement in collaborative care and behavioral activation.

Core Components of Collaborative Care

Effective delivery of collaborative care relies on five core principles [425]: patient-centered, population-based, measurement-based, evidence-based, and accountable. Using the core principles of collaborative care as a basis for determining key tasks and processes, a comprehensive list of potential technology opportunities for enhancing collaborative care has been proposed [28, 326]. An important component that drives the technology capabilities that support various clinical tasks is the patient registry[5]. A patient registry collects and organizes data about a specific population of patients to evaluate their outcomes and plays a central role in the collaboration between care managers and psychiatrists in collaborative care (Figure 2.1). A traditional patient registry in the collaborative care context typically lives within the bounds of the clinic (e.g., Excel-based registry) and is used between care managers and psychiatrists. However, to support our design opportunities, we envision a technology-enhanced registry that would receive direct input from the patients to BHPs and send documents and assignments from BHPs to patients (Figure 4.4).

A technology-enhanced collaborative care registry would connect with patient-accessible interfaces to collect patient outcomes (DO1), to capture and share documented care plans and patient history (DO4), and to enable communication between patients and providers (DO5) for patient-centered care. It can coordinate and facilitate the delivery of educational resources (DO6) and digitized evidence-based psychosocial interventions (DO3) for evidence-based care. A registry can also provide alerts for patients that have not exhibited anticipated outcome improvement to support measurement-based care, track all patients for population-based care (DO2), and monitor outcomes at a provider or an organization level for quality improvements and accountable care. In the collaborative care setting, many of these actions surrounding the registry fall on the shoulders of a BHP (typically a clinical social worker), who are tasked to work closely with the care team to coordinate and deliver comprehensive, patient-centered care. Much of the data that goes into the registry would come from the patients and the BHPs' encounters with their patients.

---

[5] https://www.sciencedirect.com/topics/medicine-and-dentistry/patient-registry

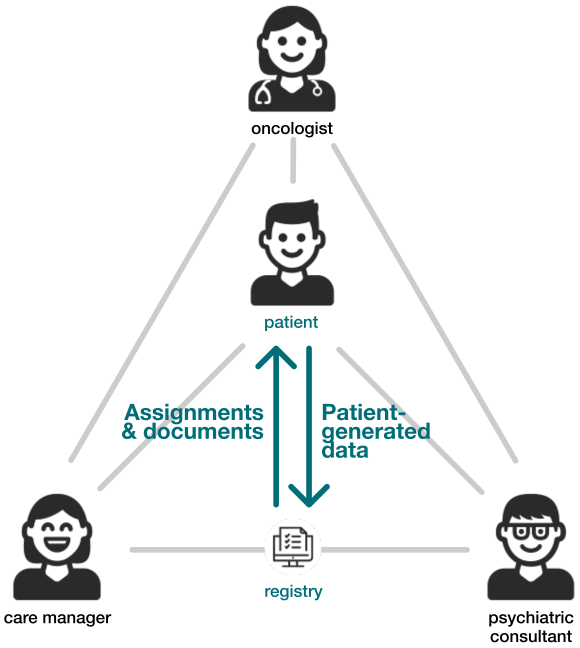


Figure 4.4: Illustration of a modified Collaborative Care Model with the technology-enhanced registry that allows direct connection between the patients and the registry. Patients can send patient-generated data to the registry, and the registry can facilitate sharing of documents and assignments from the providers.

Core Components of Behavioral Activation

At the core of behavioral activation is the realization that depression can stem from a vicious cycle of doing less of what matters which leads to a depressive mood that discourages further action. To combat the negative cycle of depression, behavioral activation encourages people to do more of what matters which leads to a positive mood and a sense of accomplishment that encourages further positive activities. Thus, activities, emotions, and their relationships are core aspects of behavioral activation [208].

Behavioral activation has been shown to be effective in primary and general care settings due to its flexible and modular delivery format [411]. A clinical guideline for behavioral activation treatment outlines the delivery of several skills to patients [33, 428]. First, clinicians are recommended to review the behavioral activation model and educate patients about the relationship between activity and mood. Then, clinicians guide patients to assess personal values and set individualized goals. Using these values and goals, patients schedule pleasant activities and track the association between mood and activities. When faced with barriers or avoidance behaviors, patients are encouraged to employ problem-solving techniques.

Of these skills, we optimized for digitizing a subset that maximizes patient-provider collaboration. We specifically target those skills that require patients to engage independently but their skill uses can be monitored and assessed by BHPs remotely to accommodate the situational burdens of cancer and to increase patient self-efficacy. These include values and goals assessment, activity scheduling, and activity/mood tracking. A technology-enhanced behavioral activation (DO3) would allow patients to use their mobile devices to receive prompts from their BHPs regarding the treatment plan and assignments (DO4). Patients can use their mobile devices to brainstorm values and goals that may drive their behavior and identify important and enjoyable activities that align with their values and goals. Patients would schedule these activities and receive reminder notifications to perform them. Patients can log their activities and assess their moods or outcomes (DO1). Patients can send their engagement and assessment data to their BHPs for monitoring (DO2). Patients can reach out to BHPs when needed and can receive feedback on their progress (DO5). Any components of behavioral activation that are not directly integrated into the system, such as problem-solving or relaxation techniques or educational materials, can be accessed through their mobile devices as separate modules (DO6).

Technology-Enhanced Collaborative Behavioral Activation Loop

In an ideal setting, a successful behavioral activation strategy would involve discussing patient values, goals, and activities during the active session with the BHPs. Outside of the session, patients would actively engage in activities and track their mood associated with those activities. At the next session, patients and BHPs would review tracked data and adjust the treatment and action plan for the patient to follow outside of the session, thus ensuring continuity in care. However, in our prior study, we found that behavioral activation in cancer settings often does not progress in this idealized, sequential manner. We found that, even though patients and BHPs may discuss values, goals, and activities during sessions, patients may experience cognitive impairments from cancer treatment side effects (i.e., “chemo brain”) that make it difficult to remember the discussions. Patients may be too overwhelmed by their disease to prioritize participating in behavioral activation. In some cases, patients cannot go to the subsequent sessions due to treatment side effects. In other cases, patients need to talk about a new issue that came up which leads to the discontinuation of the care.

An ideal behavioral activation implementation would foster a closed-loop interaction between patients and BHPs that connects one session to another but also between sessions. Guided by the core components of collaborative care and behavioral activation and the potential technology capabilities to support them, we formalize such a loop into a technology-enhanced collaborative behavioral activation loop model with actions and tasks that must be strung together to complete the loop. Figure 4.5 illustrates how each of these touch points connects to complete the collaborative behavioral activation loop. For example, during the session, patients and BHPs can record their care plan into a centralized system (Figure 4.5a). Between sessions, patients are reminded of their care plan and engage in the activities and track progress (Figure 4.5b), while BHPs monitor patients’ engagement through shared data and communication (Figure 4.5c). If patients cannot attend the next session or the next session is too far away, BHPs can send updated care plans or nudges to self-help resources. When they come back to the next session, they review the progress based on information that has already been collected (Figure 4.5d). At the end of the session, they make adjustments to the care plan as necessary and record the updated care plan, thus closing the loop (Figure 4.5a).

The model introduces four touch points to support technology. These four touch points can guide the identification of patient and provider task scenarios for which to design our system. While *Plan* and *Review* are performed during active collaboration in sessions, *Engage* is done independently by the patients, and *Monitor* is done independently by the BHPs.

Plan Before concluding the session, BHPs discuss treatment plans with their patient and sends assignments to patients through the registry to complete remote assessments, review

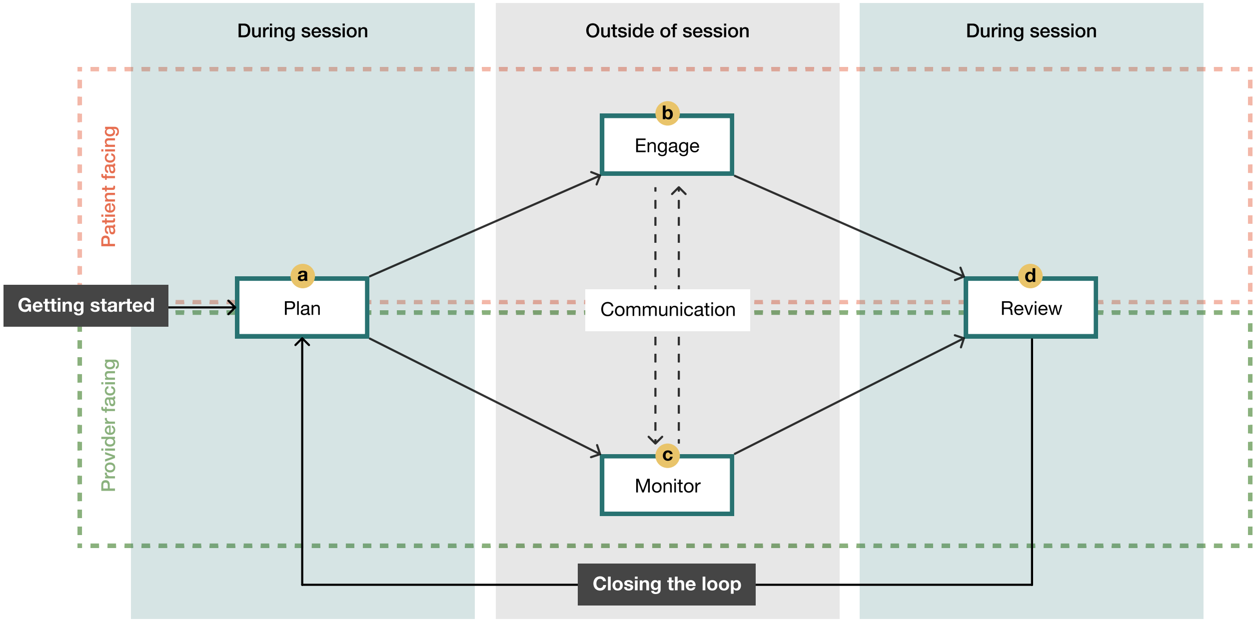


Figure 4.5: The collaborative behavioral activation loop model consists of four touch points: (a) plan, (b) engage, (c) monitor, and (d) review. During the session, patients and BHPs review the progress and make adjustments to the treatment plan. Outside of the session, patients engage in their care, and BHPs monitor patient-generated data. When they come back to the following session, they review the progress and make future plans, thus closing the loop.

educational resources, or engage in psychosocial intervention strategies. Following the session, BHPs document session output into the registry, including treatment plans, outcome measures, changes to medications, referrals, psychosocial intervention strategies used, and general session notes.

Engage Between sessions, the patient gets notified of any remote assessments that are requested by the BHPs and performs them. Patients engage in activities related to behavioral activation, including values and goals assessment, activity scheduling, and tracking activity and mood. They also review educational resources and other behavioral activation skills. Their engagement data is shared with their BHPs in real-time. Patients can reach out to BHPs for questions or urgent help.

Monitor Between sessions, BHPs use the registry to monitor engagement and assessment data from patients and review their progress to make adjustments to assignments or to reach out for flagged events or behaviors. With consulting psychiatrists, BHPs review their patient caseload to identify patients to discuss. They collaborate on the registry to determine the next steps or document recommended treatment adjustments. Before the session, BHPs review patient data in the registry, including their assessment data, prior session, case review information, and engagement history, to prepare for an upcoming session with a patient.

Review During the session, BHPs and patients review the data collected from prior assignments and troubleshoot if there were any barriers to making progress or engaging in activities. BHPs proceed with their treatment and conclude the session by discussing the action plan following the session, thus closing the loop.

4.2.2 SCOPE System Design

Using the collaborative behavioral activation loop model (Figure 4.5) introduced earlier to organize main touch points for technology support, we designed a collaborative care system called SCOPE – Supporting Collaborative Care to Optimize Psychosocial Engagement – that involves a patient-facing app and a provider-facing registry that are used together to support a collaborative experience.

The patient-facing app is built as a progressive web app that is mobile-optimized and accessible by any device and can be installed like a native app. The registry is a web-based application that providers can securely access from any web-enabled device. Both applications are backed by a centralized web service and a database that hosts patient data. Patients can only interact with their own data, while care managers and psychiatrists can interact with all patient data in their organization through the registry. Our system is designed specifically for psychosocial care and is independent of any existing EHRs such that providers outside of the psychosocial care team do not have access to the patients' sensitive data, and integrating with existing EHRs is out of scope for this deployment. Figure 4.6 illustrates the full SCOPE system capabilities with patient

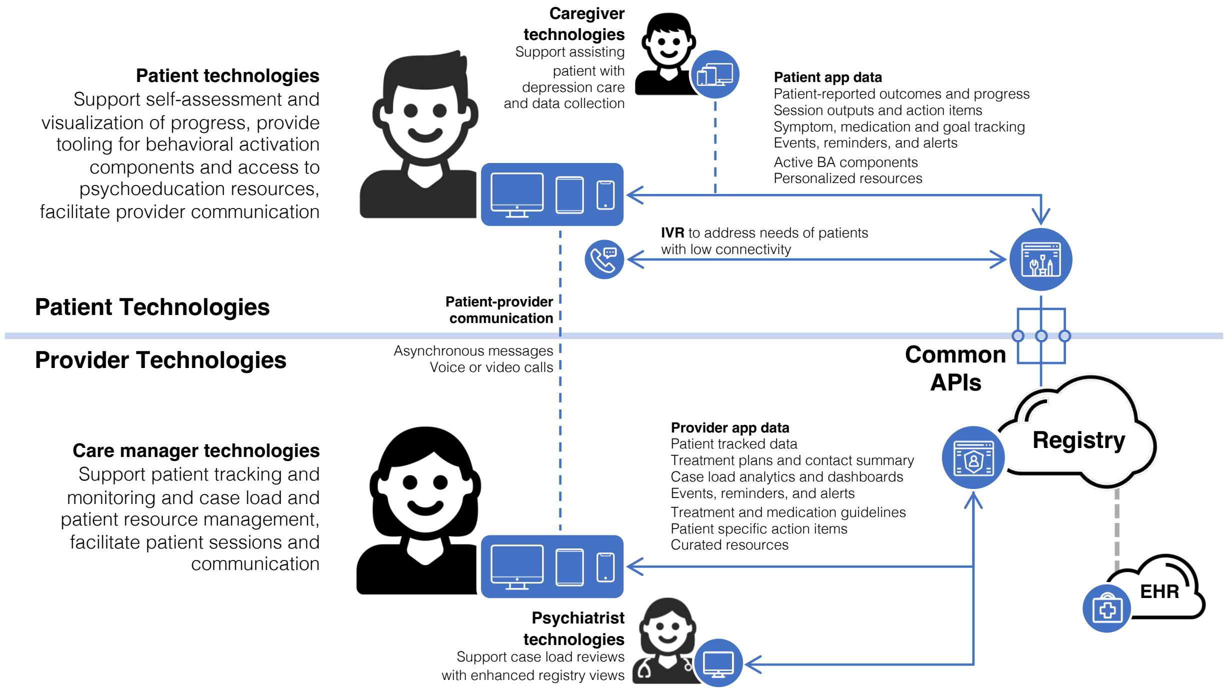


Figure 4.6: Illustration of the full SCOPE system capabilities consisting of patient and care manager technologies connected through the registry. This illustration includes system capabilities that are not yet developed, such as interactive voice response (IVR), caregiver technologies, and EHR integration.

and provider technologies. It also enumerates system capabilities that are not yet developed, such as interactive voice response (IVR), caregiver technologies, and EHR integration. Table 4.3 enumerates currently implemented SCOPE system capabilities and how each of the four touch points in the collaborative behavioral activation loop is supported by the SCOPE system.

Provider-facing registry

The SCOPE registry facilitates patient monitoring and management with integrated tools for supporting behavioral activation.

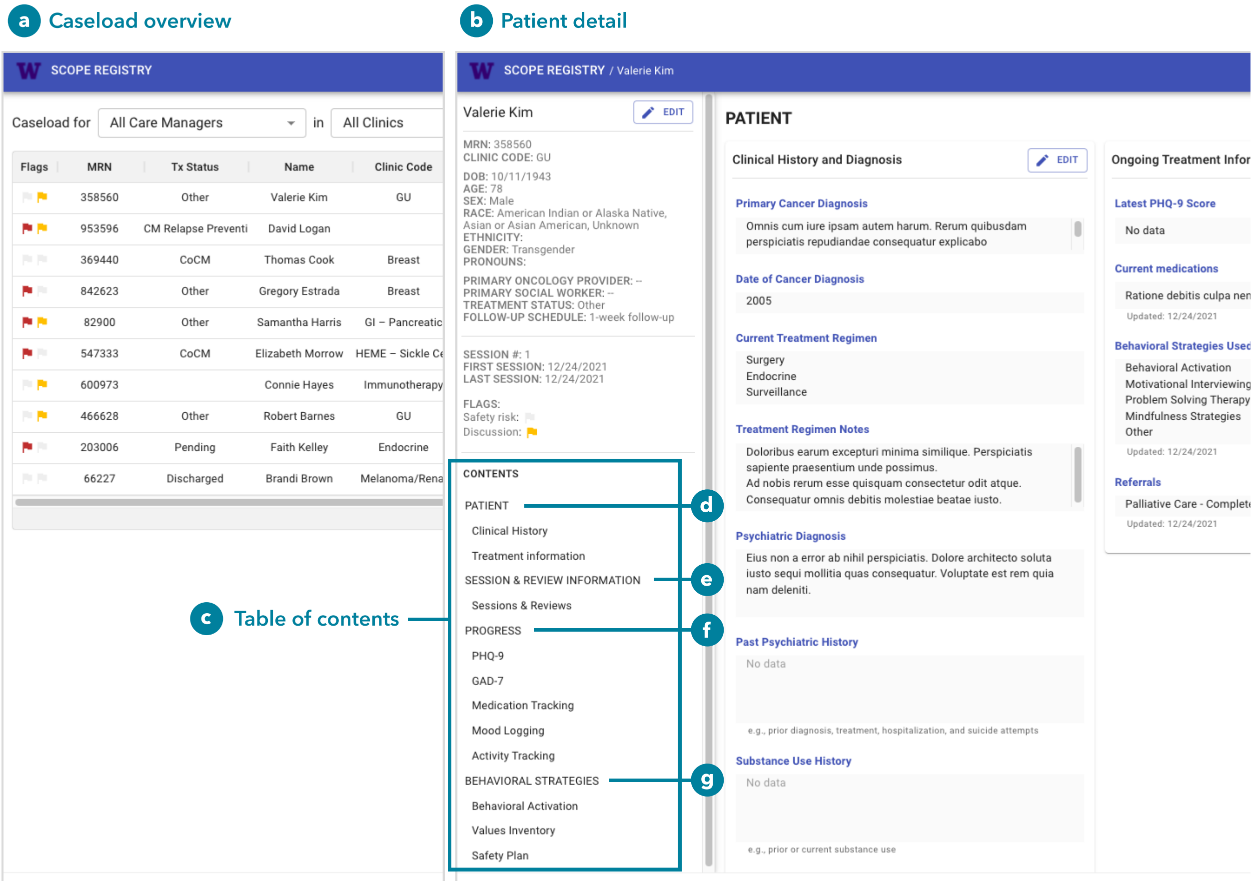
The landing page of the registry (Figure 4.7a) is an overview of the patient caseload to support patient management and prioritization. The caseload table is designed to provide a quick glimpse into patients and to help identify those that need extra attention. The caseload can be filtered

	Plan	**Engage**	**Monitor**	**Review**
Patient-facing system capabilities				
Completing depression and anxiety check in		x		
Logging mood		x		
Scheduling an activity	x	x		
Viewing and managing scheduled activities		x		x
Logging whether or not you have completed a scheduled activity		x		
Viewing shared documents and learning resources	x	x		
Completing values and activities inventory	x	x		
Viewing depression and anxiety check-ins		x		x
Viewing logged activity data		x		x
Viewing logged mood data		x		x
Completing safety plan		x		
Viewing crisis help resources		x		
Sending patient data to care team		x		
Provider-facing system capabilities				
Assigning assessments (e.g., PHQ-9, GAD-7) to patients	x			
Assigning values inventory	x			
Assigning safety plan	x			
Accessing behavioral activation resources	x			
Viewing and editing patient history and information	x		x	x
Viewing and editing session information	x		x	x
Viewing completed behavioral activation components			x	x
Viewing patient progress on depression and anxiety			x	x
Viewing patient progress on mood			x	x
Viewing patient progress on scheduled activities			x	x
View patient values and activities inventory			x	x
View patient safety plan			x	x
Flagging patients for safety or discussion			x	
Viewing and interacting with the list of patients in the caseload table			x	
Viewing and editing caseload review information			x	

Table 4.3: List of SCOPE system capabilities and how they support the four touch points of the collaborative behavioral activation loop.

by a care manager or a clinic within that organization and allows sorting across columns. The table provides summary information such as the first and last session, the last case review, total number of sessions. It tracks patient-level flags (e.g., safety risk, need discussion) and when the follow-up is due to identify patients who have not been seen in a while. It provides patient outcome assessment results when they were administered, and the changes in outcomes to help identify patients that have not improved and need attention. The caseload table also acts as an entry point to the individual patient overview.

The patient overview page (Figure 4.7b) consists of patient profile and history, session and case



Caseload for All Care Managers in All Clinics

Flags	MRN	Tx Status	Name	Clinic Code
⚐ ⚑	358560	Other	Valerie Kim	GU
🚩 ⚑	953596	CM Relapse Preventi	David Logan	
⚐ ⚐	369440	CoCM	Thomas Cook	Breast
🚩 ⚐	842623	Other	Gregory Estrada	Breast
🚩 ⚑	82900	Other	Samantha Harris	GI – Pancreatic
🚩 ⚐	547333	CoCM	Elizabeth Morrow	HEME – Sickle Ce
⚐ ⚑	600973		Connie Hayes	Immunotherapy
⚐ ⚑	466628	Other	Robert Barnes	GU
🚩 ⚐	203006	Pending	Faith Kelley	Endocrine
⚐ ⚐	66227	Discharged	Brandi Brown	Melanoma/Rena

Valerie Kim

MRN: 358560
CLINIC CODE: GU

DOB: 10/11/1943
AGE: 78
SEX: Male
RACE: American Indian or Alaska Native, Asian or Asian American, Unknown
ETHNICITY:
GENDER: Transgender
PRONOUNS:

PRIMARY ONCOLOGY PROVIDER: --
PRIMARY SOCIAL WORKER: --
TREATMENT STATUS: Other
FOLLOW-UP SCHEDULE: 1-week follow-up

SESSION #: 1
FIRST SESSION: 12/24/2021
LAST SESSION: 12/24/2021

FLAGS:
Safety risk: ⚐
Discussion: ⚑

CONTENTS

PATIENT
  Clinical History
  Treatment information
SESSION & REVIEW INFORMATION
  Sessions & Reviews
PROGRESS
  PHQ-9
  GAD-7
  Medication Tracking
  Mood Logging
  Activity Tracking
BEHAVIORAL STRATEGIES
  Behavioral Activation
  Values Inventory
  Safety Plan

PATIENT

Clinical History and Diagnosis

Primary Cancer Diagnosis
Omnis cum iure ipsam autem harum. Rerum quibusdam perspiciatis repudiandae consequatur explicabo

Date of Cancer Diagnosis
2005

Current Treatment Regimen
Surgery
Endocrine
Surveillance

Treatment Regimen Notes
Doloribus earum excepturi minima similique. Perspiciatis sapiente praesentium unde possimus.
Ad nobis rerum esse quisquam consectetur odit atque.
Consequatur omnis debitis molestiae beatae iusto.

Psychiatric Diagnosis
Eius non a error ab nihil perspiciatis. Dolore architecto soluta iusto sequi mollitia quas consequatur. Voluptate est rem quia nam deleniti.

Past Psychiatric History
No data
e.g., prior diagnosis, treatment, hospitalization, and suicide attempts

Substance Use History
No data
e.g., prior or current substance use

Ongoing Treatment Infor

Latest PHQ-9 Score
No data

Current medications
Ratione debitis culpa nen
Updated: 12/24/2021

Behavioral Strategies Used
Behavioral Activation
Motivational Interviewing
Problem Solving Therapy
Mindfulness Strategies
Other
Updated: 12/24/2021

Referrals
Palliative Care - Complete
Updated: 12/24/2021

Figure 4.7: Screenshots from the SCOPE registry. The registry supports patient caseload overview (a) to identify high-risk patients. For each patient in the registry (b), providers can enter patient history (d), add session and case review notes (e), monitor patient progress (f), and review behavioral strategies and resources (g). Providers can jump to various sections from the table of contents (c).

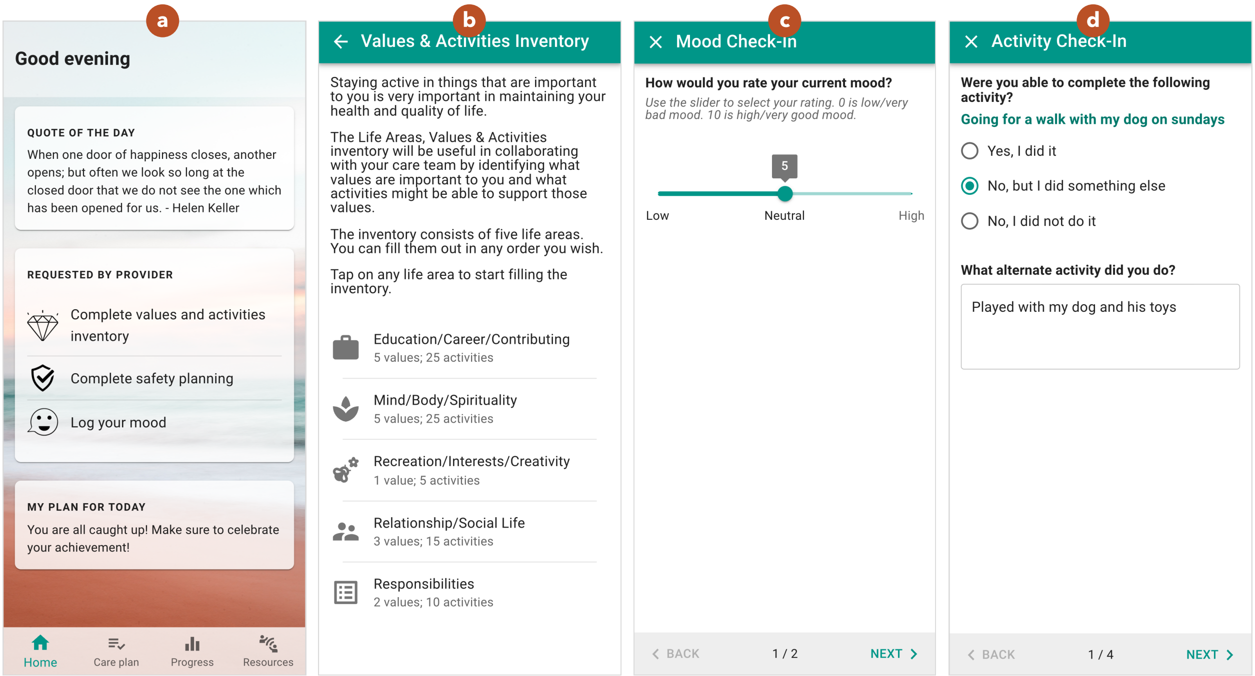
review information, patient assessment results and progress, and behavioral strategies employed to treat the patient. The patient profile panel provides the medical record number (MRN), clinic code, demographics (e.g., age, gender), and care team information (e.g., primary oncologist, social worker) that are affixed to the left side for easy access. It also includes flags for safety risks and discussion at case review meetings. Providers can edit patients' profile information and toggle these flags. Below the patient profile is a navigation panel (Figure 4.7c) that allows quick access to different sections of the patient overview.

The patient section (Figure 4.7d) includes clinical history and diagnosis for both cancer and mental health which is available for editing. The summary of ongoing treatment information is automatically pulled from an existing session and assessment data and placed alongside clinical history so that providers can get a patient's background information from a single place.

Session and review information section (Figure 4.7e) displays a table enumerating the history of all patient encounters and case review discussions. Providers can add a session or a case review to bring up a modal dialog to enter relevant notes. Inside a session information modal dialog, providers can track session type, session duration, medications, behavioral strategies used during the session, referrals made during the session and their status, recommendations, action items, and session notes. Inside a case review modal dialog, providers can track the consulting psychiatrist, any changes made to medications or behavioral strategies, referrals, recommendations, action items, and review notes.

Progress section (Figure 4.7f) displays patient-generated data, such as depression assessment (PHQ-9), anxiety assessment (GAD-7), mood logging, and activity tracking. As patients enter assessment data through their apps, each assessment displays a table enumerating each entry as well as a visualization of trends in reported outcomes over time. The visualization supports showing the total score for assessments or the individual items. Providers can toggle each item on or off to view a subset of items at a time. Providers can assign an assessment for a specific frequency and day of the week based on the needs of their patients, and the assigned assessments will show on patients' apps as requested. Providers can also enter assessment data if the assessment was conducted in the clinic or over telehealth. Depending on the level of granularity providers need for tracking assessment data, they can enter one total score or each score of individual items in the scale.

Behavioral strategies section (Figure 4.7g) includes information about behavioral activation treatment and patient-entered values inventory and safety plan. The behavioral activation panel enumerates each component of behavioral activation and tracks when it was last completed in sessions. It also lists relevant forms and worksheets that providers can review with patients. The same resources are also available on the patient-facing app so that they can access them at any time. Providers can assign values and activities inventory to patients, which will show up as



Good evening

QUOTE OF THE DAY

When one door of happiness closes, another opens; but often we look so long at the closed door that we do not see the one which has been opened for us. - Helen Keller

REQUESTED BY PROVIDER

Complete values and activities inventory

Complete safety planning

Log your mood

MY PLAN FOR TODAY

You are all caught up! Make sure to celebrate your achievement!

Values & Activities Inventory

Staying active in things that are important to you is very important in maintaining your health and quality of life.

The Life Areas, Values & Activities inventory will be useful in collaborating with your care team by identifying what values are important to you and what activities might be able to support those values.

The inventory consists of five life areas. You can fill them out in any order you wish.

Tap on any life area to start filling the inventory.

Education/Career/Contributing
5 values; 25 activities

Mind/Body/Spirituality
5 values; 25 activities

Recreation/Interests/Creativity
1 value; 5 activities

Relationship/Social Life
3 values; 15 activities

Responsibilities
2 values; 10 activities

Mood Check-In

How would you rate your current mood?

Use the slider to select your rating. 0 is low/very bad mood. 10 is high/very good mood.

Activity Check-In

Were you able to complete the following activity?

Going for a walk with my dog on sundays

○ Yes, I did it
● No, but I did something else
○ No, I did not do it

What alternate activity did you do?

Played with my dog and his toys

Figure 4.8: Screenshots from the SCOPE behavioral activation app. The patient app supports the core components of behavioral activation. From left to right, each screenshot shows (a) provider-assigned and patient-scheduled action items, (b) values and activities inventory, (c) activity logging, and (d) mood logging.

requests on the patient's app. When the patient fills out the inventory from their app, the providers can see the list of activities and their associated enjoyment, importance, life area, and value. For patients with safety concerns, providers can assign a safety plan to be completed. The patient's submitted safety plan will then be visible to the provider for review.

Patient-facing behavioral activation app

The SCOPE patient app facilitates behavioral activation and remote self-assessments. It also provides access to a personalized safety plan and treatment resources.

The home page of the patient-facing app (Figure 4.8a) includes an inspirational quote of the day, tasks requested by the provider, and activities planned for the day. The tasks requested by the provider could include values and activities inventory, safety plan, depression check-in, anxiety check-in, or mood logging. Activities planned for the day will be populated based on the activities that patients schedule for behavioral activation.

When a patient accesses their values and activities inventory (Figure 4.8b), they choose from five life areas: Education/Career/Contributing, Mind/Body/Spirituality, Recreation/Interests/Creativity, Relationship/Social Life, and Responsibilities. In each of the life areas, patients can add a list of values associated with the life area. For each value, patients can add a list of activities that are associated with the value. Patients can access the values and activities inventory at any time to modify the list of values and activities.

When a patient accesses their safety plan, they answer a series of paginated prompts including reasons for living, warning signs, coping strategies, social distractions and support, professional support, and a safe environment.

When a patient enters a depression or anxiety check-in flow, each question of the scale is displayed in a paginated manner. After all questions are answered, patients can see their final score, including a table that helps interpret the score. For example, a score range of 5 to 9 in PHQ-9 would display "Your symptoms are relatively mild. Keep up the good work as you and your clinical social worker collaborate to decrease them even more!" If the patient responds positively to the suicidality question of the PHQ-9, the app would display crisis resources.

In the mood logging flow (Figure 4.8c), patients are asked to rate their current mood from a scale of 0 (Low) to 10 (High). Then, they are asked to write any notes relevant to the mood that can be shared with the clinical social worker.

Patients can mark activity completion by clicking on an outstanding activity (Figure 4.8d). If the patient marks the activity as 'Yes, I did it', they are asked to rate the level of accomplishment and pleasure in performing the activity. If the patient marks the activity as 'Yes, but I did something else', they are also asked to provide the alternative activity they performed. If they answer 'No, I didn't do it', they are prompted to write thoughts on barriers to doing that activity.

The patient app has three additional pages accessible via the bottom application bar: activities, progress, and tools.

The activities page displays a list of scheduled activities that the patient has configured in the activity view. In the calendar view, patients can navigate to a different day on a calendar to see which activities are scheduled for that day. Patients can schedule a new activity from this page. Patients can choose an activity from their inventory to schedule, or they can schedule a brand-new activity. Patients can set a date and time for the activity and configure a repeat and/or a reminder for the activity. Once scheduled, the scheduled activity appears in the list where the patient can delete or deactivate it as necessary.

The progress page enables access to tracked data for depression, anxiety, activity, and mood. Tracked data is displayed in a table, and patients can click into each row to see full submitted data, including individual scores or notes.

The tools page contains links to the values and activities inventory, a library of shared documents and learning resources, a safety plan, and crisis help. The library of resources contains identical forms and worksheets to the ones shown in the provider-facing registry. The tools page also includes information about the research team and a button to log out of the app.

4.2.3 Study Design and Method

Our research goal is to design a technology-enhanced collaborative behavioral activation system and to deploy the technology in a randomized clinical trial (RCT) that compares the efficacy of technology-enhanced care against that of usual care. The SCOPE system (Section 4.2.2) described earlier was a result of three design iteration studies and a pilot study. To evaluate if the SCOPE system supports collaboration and engagement as intended, we analyze the data from these four user studies. In this section, we describe the design process and the analysis method.

Design process

Our design process involved (1) three major iterations with stakeholders using wireframes and interactive prototypes and (2) one pilot study with a deployed system in a real clinical setting. Each iteration required a rapid turnaround between design and evaluation to prepare for the RCT deployment. In making design decisions, we weighed inputs from our expert collaborators who are practicing clinicians in collaborative care settings to prioritize various technology capabilities. We also weighed inputs from our engineers to determine the feasibility and cost associated with developing various technology capabilities. With each iteration, we leveraged our design and prototype as a probe and a data collection mechanism to gather user requirements [50, 194]. In particular, we sought (1) to obtain technology design requirements that meet our stakeholder needs, (2) to field test our prototype, and (3) to characterize potential challenges and tensions that would hinder stakeholder engagement in a real-world deployment of the system. Our design and technology probe provided a grounding for facilitating discussion with potential users of the system around their challenges and needs. Our design process follows primarily a convergent approach from iteration to iteration because the ultimate aim of our project is to develop a technology solution that can be deployed and tested in an RCT.

Each iteration was followed by an interview study with stakeholders. Each study focused on a specific subset of scenarios that required stakeholder inputs, and subsequent user studies modified, replaced, or expanded on prior scenarios to better support findings from prior studies or to investigate new areas. We captured the implementation outcomes [439] and the System Usability Scale (SUS) [237] to gauge whether our design is on the right track. To facilitate the rapid design process, the interviewers synthesized the feedback and captured design action items from one interview to make modifications to the existing design for the next interview. Therefore, none of the interview data was formally analyzed during the design process. All interviews were conducted and recorded over Zoom and transcribed using third-party software (Otter.ai).

Design Iteration 1 The first design iteration involved exploring patient app functionalities around behavioral activation support and self-assessment.

To conduct the evaluation, we recruited 6 patients from cancer clinics participating in our RCT. Of 6 patients interviewed, 5 reported as Caucasian/White and 1 as multiracial. Four patients reported as female and 2 as male. Patient ages ranged from 21 to 77 years old ($\bar{x} = 57.2$, $\sigma = 20.2$).

To evaluate the design, we identified 6 scenarios: mood logging, activity scheduling, logging complete activities, logging incomplete activities, remote self-assessment, and sending a message to the provider. We designed a 90-minute, semi-structured interview protocol around these 6 identified scenarios. For each scenario, we walked patients through the app screen by screen on how a hypothetical patient would use the app to complete that scenario. Following each scenario, we asked patients to rate implementation outcomes and asked open-ended questions. We assessed the implementation outcomes using modified, single-item versions of the Acceptability of Intervention Measure (AIM), Intervention Appropriateness Measure (IAM), and Feasibility of Intervention Measure (FIM) on a scale from 1 (Strongly disagree) to 5 (Strongly agree) [439]. We then asked open-ended questions about their likes, challenges, and desired changes. After reviewing all scenarios, we asked patients to assess the usability of the design prototypes using the System Usability Scale (SUS) [237]. If there was remaining time, we asked patients to provide comments about the scenarios, potential design ideas for the home page and reminders, and other considerations for patients like themselves in designing the app.

On average, our design received 92.1 (σ=7.31) out of 100 on the SUS scale and 4.59 (σ=0.66) out of 5 on the implementation outcome scales.

Design Iteration 2 The second design iteration involved building out the near full set of wireframes for both the patient app and the registry according to the tasks identified by the collaborative behavioral activation loop.

We recruited 8 patients to evaluate the patient app design and 6 social workers to evaluate the registry from the cancer clinics participating in our RCT. Of 8 patients interviewed, 4 reported as Caucasian/White, 2 as African American/Black, 1 as American Indian, and 1 as multiracial. Six patients reported as female and 2 as male. Patient ages ranged from 31 to 78 years old ($\bar{x} = 58.9$, $\sigma = 13.3$). Of the 6 social workers interviewed, all reported as Caucasian/White. Five reported as female and 1 as male. Social worker ages ranged from 34 to 58 years old ($\bar{x} = 47.2$, $\sigma = 11.0$), and 1 did not report their age.

This design iteration focused on achieving high fidelity to evidence-based care practices. In our patient app design, we aligned our design to follow the behavioral activation practices that will be implemented in the RCT. For example, we included a full values and activities inventory flow in the app. Unlike our previous design that used descriptive language for mood scales (i.e., awful, bad, okay, good, great), we used a standard mood tracking scale (i.e., low, neutral, high from 0 to 10) that was used in practice. We also removed the communication scenario from our previous iteration for two reasons: (1) our expert collaborators were concerned about the potential for added provider burden [109], and (2) existing EHRs provide a way for patients and providers to communicate. Our registry design included viewing caseload overview, creating a new patient record, adding session information, reviewing patient progress, preparing for a case review, and presenting the patient information during a case review.

The interview procedure for the second study was nearly identical to the first study, with each scenario walkthrough followed by implementation outcome ratings and open-ended questions for likes, challenges, and desired changes. In this study, we included a Net Promoter Score (NPS) [329] to gauge the interest in stakeholders for sharing the technology with others. We also asked social workers to design a new organization of patient information that would help them present the patient information during the case review. We used an online interactive whiteboard (Miro[6]) to facilitate the discussion where each patient information unit could be freely moved and reorganized.

On average, our design received 83.21 (σ=13.67) out of 100 on the SUS scale, 4.49 (σ=0.77) out of 5 on the implementation outcome scales, and 9 (σ=1.22) out of 10 on the NPS.

Design Iteration 3 The third design iteration involved developing a functional prototype that stakeholders can test themselves. The primary audience was the providers who would use the system in actual care practice, and our goal was to help them evaluate how the patient app and the registry can be integrated into their workflow. We recruited 6 social workers from the same two cancer clinics. All 6 social workers identified as Caucasian/White, and 1 reported as being Hispanic. Four reported as being female and two as male. Social worker ages ranged from 32 to 58 years old ($\bar{x} = 44.1$, $\sigma = 10.1$).

In the functional prototype for the registry, we implemented most functionalities, including an interactive visualization for patient-submitted assessment data. The prototype also demonstrated how data from different sections of the registry are integrated to provide a holistic view of the

---

[6] https://miro.com/

patient and the treatment. The final design included the reorganization of patient information as the previous study had suggested. In the functional prototype for the patient app, we implemented most screens for BHPs to simulate the patient experience.

The interview procedure involved sharing a link to the prototype that participants could access from their machines. We asked participants to share their screens and to speak their thoughts out loud as they interacted with the prototypes. We did not walk through specific scenarios as in the first two studies, but we guided participants to different parts of the apps to get their impressions. The participants first experienced the registry and then explored the patient app. They commented on how they would expect patient-provider engagement to happen. We did not ask any questions that captured quantitative data.

Pilot Study We made further iterations on the design based on the feedback from the previous three studies. We developed a fully functional registry and a patient app that is connected to the back-end service and database. We deployed the system to cancer clinics and enrolled participants in the program in preparation for the RCT.

After several months of the pilot enrollment, we conducted a semi-structured interview study to get initial feedback and user requirements from actual deployment and usage of the system in a clinical setting. Unfortunately, 5 of our 6 pilot patients refused to be interviewed or had difficulties in scheduling interviews. Engaging patients with cancer to engage in research activities is often challenging due to their already over-burdened lives [378]. We only recruited 1 patient but were able to interview all 3 social workers and 1 psychiatrist enrolled in the pilot study. All participants reported as Caucasian/White and female. Their age ranged from 37 to 62 years old ($\bar{x} = 50.2$, $\sigma = 10.1$). We do not report demographics per stakeholder type due to the increased risk of identification for the small participant size.

During the interview, we first asked to describe how participants used the system. Then we asked how the system supported collaboration and coordination of care amongst stakeholders, how the system supported data collection and monitoring, how the system supported shared decision-making and adjustments to care, and how the system supported adherence to guideline-level care.

Analysis

Our system design process required a rapid turnaround between design and evaluation. After the four rounds of studies, we aimed to systematically analyze the interview data to evaluate how the current design does or does not support various stakeholder tasks identified in our collaborative behavioral activation loop. We also aimed to understand how stakeholders would integrate the system into their collaborative and independent care experiences.

We first combined interview transcripts from all four studies to create a single dataset. We used inductive thematic analysis [57] to conduct a bottom-up approach to coding the interviews. Wherever applicable, each code was associated with a design iteration, a stakeholder type, and a scenario that was discussed. Each code was subsequently merged for duplication and grouped under similar themes. Each theme was then categorized into the four touch points – *Plan*, *Engage*, *Monitor*, and *Review* – identified in our collaborative behavioral activation loop model. We included two additional categories – *Getting Started* and *Closing the Loop* – to describe feedback surrounding the initial and continued engagement with the system in the care practices. As a result, each category was associated with a set of themes that represented stakeholder needs that were supported or not supported and themes that represented how each stakeholder type envisioned the system being integrated into their care experiences.

Participant Anonymization Our studies involved interviewing social workers who were employed at the cancer clinics involved in our RCT. Some of the clinics only had a few social workers that were responsible for caring for hundreds of patients. Our studies, therefore, involved interviewing the same social worker in multiple design iterations. Our study also included one of only a few psychiatrists who have shared consulting duties across cancer clinics. To preserve the anonymity of our participants, we do not differentiate social workers from psychiatrists. We also do not differentiate which social worker participated in which design iteration nor do we reveal which social worker participated multiple times in our studies. In our findings, we refer to all 16 interviews with social workers and psychiatrists as BPH1 through BPH16. Since there were no repeat patients in our studies, we refer to all 15 patient interviews as Pt1 through Pt15.

4.2.4 Findings

We present our findings organized by the four touch points – *Plan*, *Engage*, *Monitor*, and *Review* – identified in our collaborative behavioral activation loop model and two categories – *Getting Started* and *Closing the Loop* – that describes how the system could be integrated into the care practices.

Getting started: integrating the system in patient-provider collaboration

Can the app shed a light on depression?

> *"Because if I find myself spiraling down, and I'm having a hard time stepping off that merry go round that's going down. Because this is another tool that I would use to help myself. (Pt11)"*

Despite the app being perceived as more work, patients thought that the app would help them recognize their depression. When they feel low, they can engage in an activity *"to help [them] get out of that quicker than what [they] normally would. (Pt9)"* One patient thought that the app could help them to be more transparent about their depression: *"If I'm honest about using that there would be less places that I could hide, you know, hide information from myself or from them that could affect my treatment, my care, my well being? [...] And the app would definitely give information so that I would, I would, I would be stepping out of isolation. (Pt7)"* Some patients explained that there may be certain situations when the app may be more appropriate. One patient mentioned that the app would not be appealing if they are in severe depression: *"I've even been hospitalized. And this would not be attractive. But since I am healthy now this is appealing to me. But when I was at my most depressed it would not have been. (Pt1)"* Another patient wanted to use the app after the chemotherapy: *"Maybe when chemo is over and you're just trying to heal. This is so much helpful. And when? Because I remember like, after right after surgery and stuff, no matter how hard I try, but* *maybe this can help. I don't know, because I never tried anything like this. You know, it's like, you're just in this disease mood. (Pt4)"*

Patients need a clear justification for using the app.

> *"If my doctor said, you know, I think this is what you need, I would certainly give it my best shot. (Pt5)"*

There were mixed reactions to being motivated to use the app by patients, as some saw clear value in understanding the association between mood and activity and some were not interested or were hesitant in engaging with the app. A few patients thought that the activities requested of the app were a lot of work: *"It is making me feel [...] that here's yet another task that I have to do and I'm tired. (Pt10)"* And 7 BHPs were concerned that the app would add more work to already overburdened patients. However, 9 patients expressed, and 2 BHPs agreed, that the willingness to use the app to track their mood or activities depended on a clear justification and a need from their BHPs and a commitment from the patients. Some patients were unsure how the app would help their depression, and how the BHPs would use the data to improve their symptoms: *"Who is accessing this information and, you know, my care teams accessing this information, how are they going to use it? To help me? (Pt7)"* Patients also wanted assurance that the care team would be committed to taking advantage of the app in their treatment: *"It's going to have to be something that my care team is familiar with. Otherwise, you know, the data capture and the usefulness in in that face to face interaction with your team is going to be limited if they're if they're not willing to buy into the this tool. (Pt8)"*

It takes two to collaborate.

> *"I am concerned about a lack of social work enthusiasm for the registry and the app. (BHP16)"*

Although many of the BHPs have reported using the registry for case reviews with psychiatrists, there has been unenthusiastic adoption of the system as a whole. One BHP explained that there is no need to use the registry if their patients are not using the app (BHP14). Another BHP explained that the use of the system really depends on the situation that their patients are in and their motivation level: *"I have had a really hard time. Okay, not because of the app or the registry, but because of where my patients are at. Okay, my patients are really ill. They've been in and out of the hospital. They've had some cognitive changes, and they just weren't getting on it. [...] I think it has to do with their lives. Yeah. And then I've got one that is so fully engaged. It's and I don't have to do much. She just kind of does that she loves it. She um, she writes notes, like when she does a PHQ. She tells you kind of how she's feeling when she did that. And anyway, so she doesn't need much direction. (BHP13)"*

In some cases, a challenge on one side of the collaboration breaks the whole cycle. For example, patients may not be set up in the system to engage: *"Because I didn't have... I hadn't launched the PHQ-9, the GED7, she couldn't fill one out. (BHP14)"* Or, the patients are properly set up and sharing data, but the BHPs have trouble getting into the system: *"there's one social worker right now, who has not gotten their sign in for the registry. And so their patient is using the app data is coming in. But that social worker can't and then isn't looking at what's in there. I will say it is. It can be frustrating. And then it's a less effective use of time if the social workers haven't reviewed what the patient's doing in the registry via the app, before I review it with them. (BHP16)"* But once the connection is established, the BHPs are very excited to see their patients' data: *"There's a lot more information, you know, because she does her mood about every two to three days, and then a PHQ every two weeks, so it's just regular. Yeah. Always information. So it's kind of exciting for me to look at the registry and say, oh, yeah, she's been in there. So I'm gonna look and see what she did. (BHP13)"*

Engage: improving self-care between sessions

The app augments existing therapy and life activities.

> *"Only in that if they were in fact, utilizing this tool, it would give us a another point of reference in, in our in my care. Yeah. And how? Yeah, give us another point of reference.*

> *(Pt8)"*

Many patients thought the app, especially the behavioral activation components in the app, would help guide their therapy beyond just *"talking (Pt7)"* Once completed, one patient thought that having access to the values and activities inventory summary would help provide *"direction (Pt10)"* in life and bring more *"clarity (Pt10)"* for the BHP and the patient. Another patient mentioned that the inventory would set them up *"for what's like success and all different types or areas of [their] life. (Pt9)"* Unlike a typical calendar that organizes activities throughout the day, the activities with value alignment were meaningful in a different way: *"I can think of why this could be this could be better, more, there's... there's more to it than just putting a task in my Google calendar because it's connecting with my values. (Pt7)"* The app also provided a different kind of interaction than traditionally scheduled therapy: *"I think it's a lot more interactive and a lot more a minute by minute rather than, I mean, literally minute by minute, but it's more than just Oh guys, I have a scheduled time to talk to her. That, you know, then I blast out everything I can think of at that time, really manage what I'm doing ongoing now. It really gets it off my chest. Whereas this is something that would be more like maybe she can make suggestions. (Pt9)"*

Beyond behavioral activation components, six patients mentioned that personalizing the app would draw them to use the app more. Several patients wanted a calming feeling associated with the app through calming backgrounds, music, or animations because *"[they] need help, (Pt4)"* so anything peaceful would help distract them from feeling bad (Pt1). Personalization also helps make the app theirs, as one patient said: *"Yeah, this is my page. This is... I do this. And I can do this. (Pt15)"*

Patients want flexibility in what to track.

> *"What are you like? [...] some people are numbers people, maybe they want to pick the number chart. I'm a five out of 10. Or, and some people are artistically driven. Maybe they want emojis or something like to pick out, you know, up to 10 emojis that they feel encapsulates their mood? (Pt15)"*

Although the primary target for tracking was mood and activities to support behavioral activation, many patients wanted to track other symptoms that might be relevant, such as pain (Pt3), headaches, tinnitus, nausea (Pt8), social activities (Pt7), gratitude or sleep (Pt12). Many patients had prior experience or intuition that these other symptoms affect their mood and wanted an objective way to validate what they are going through: *"I would somewhere be noting that okay, this is day one of a chemo cycle. Yeah, this is day two of the chemo cycle, because those things affect my physical state, which then affects my mood. And that is a big part of the tracking for me, okay, I know, with these chemos, on this day of the cycle, I'm likely to have these physical symptoms, and I can then it eases things for me mentally, emotionally to know, oh, this is what's going on. And I know it's gonna pass in this kind of a timeframe. And so that can definitely help my my mental state. (Pt8)"*

Different patients wanted a different way of tracking their moods or activities. While one patient preferred to describe their mood with adjectives such as *"irritable"* or *"impatient (Pt10)"*, another patient wanted the simplest logging as possible: *"the conflict could be you know, having to stop and think do I have to type in words. (Pt3)"* But many agreed that having detailed notes was helpful for their care: *"The fruit is where people are describing what is actually going on. (Pt7)"* For example, having notes can be helpful for the clinician to know if the inability to do the activity was *"out of the patient's control, versus a mood related issue. (Pt8)"* If a patient failed to do an activity, patients could dig into *"the issues around why [they] couldn't do [...] what [they] committed to do (Pt5)"* and *"decide if that's something that [they] want to do again. (Pt9)"*

Patients see values in reviewing their own data.

> *"I have a seriously sat down and catalogued and given structure to the things when I'm feeling down. So in spite of rehashing being being an evil thing, if I don't go back at some point, and try to analyze it, I'll never fix it. (Pt14)"*

Although the app was primarily designed to be used in collaboration with the providers, many patients found value in engaging in self-reflection of their own data. Many patients believed that the tracked data and visualizations will be *"informative (Pt11)"* and help them understand how *"all* *the stuff that [they're] dealing with (Pt9)"* is affecting their days. Because patients can have a track record of the activities they have not done, they will realize that they may be *"taking away the chance of [them] feeling better (Pt4)"* that will prompt them to do what is necessary to improve themselves.

Most importantly, patients thought that the tracked data gives them a different perspective: *"When you see it down on paper, or see it's somehow in a format, other than just running it through your brain? You get another? Another viewpoint? Yeah. You know, and it's, it speaks to another part of me. (Pt11)"* The data can also help patients avoid being in a false negative loop: *"For having actual data for perspective versus the mind that can lie to you at times over mood especially. (Pt8)"* However, one patient warned that knowing too much about the questions or the data could tempt them to *"skew it a little bit. (Pt12)"* Others deferred to the BHPs to help interpret the data: *"They're the expert. They might see something that I'm not... I don't recognize yet. (Pt9)"*

Monitor: improving clinical care between sessions

Registry helps provide structure around treatment and review.

> *"As the counselor, it has helped my structure of my counseling with having the library [of BA resources] having the PHQ and the GAD and seeing the progress and having the the measurements, that's what it is having the measurements visually, so that I can be tracking more over the continuum of how they're doing. (BHP15)"*

As our prior study revealed, BHPs supporting patients with cancer and depression often make frequent adjustments to their treatment plan. With the registry outlining the core components of behavioral activation as a checklist, one BHP felt that the program is more *"rigid"*, when in reality, treatment plans are not executed in a *"consecutive order, depending on what's happening with the patient. (BHP2)"* Another BHP felt that it helps structure their counseling to go beyond talk therapy: *"The positives, is having the it's like a new language. So having a new language on what behavioral activation is, and how you track it, and how you track it, your mood with it. For some of the patients, that seems to be really clicking well, and spurring them on to continue to do* *the behavioral activation. So I think having the tracking and the scales of the mood is a different language tool than us just sitting in the counseling room chatting. (BHP15)"* Access to the library of behavioral activation resources and tracking which components of behavioral activation can be used as a tool to *"fall back on"* and remind them of strategies that they *"have not inquired about (BHP1)"*, and they have been used help *"jog [their] thought of [are they] using everything available? (BHP15)"* However, one BHP felt that the system is not helpful if behavioral activation is not the main treatment strategy used with the patient (BHP14).

Beyond the session, BHPs thought that the structure of the registry and input fields helped remind them what information is relevant to record for patients, compared to their previous practice of using *"blank notes section (BHP3)"* in the EHRs, *"hunting and pecking to find information (BHP2)"* in the Excel registry , or not capturing necessary information for caseload review (BHP6). It was especially helpful for preparing patient records to review with psychiatrists: *"I have not asked about substance use or sometimes I've written down medication recommendation and a plan and the case console that I've I've the piece of paper, you know, maybe he's been misplaced or something like that. And so now it's really nice that it's an electronic electronic format because all we have right now for case console is just to say that we did it Yeah, that's it. There's no like, here's a friendly reminder. (BHP1)"* The organization of patient information in the registry was *"more in favor of it flowing in a reading format, based off of a typical presentation format that [they] give. (BHP1)"* However, the lack of integration with EHR meant that the BHPs do not get a full picture of the patient. One BHP mentioned about the benefit of EHR: *"I can see other people's inputs. So I'm not, it's not just me social work. Right, so I can integrate. So a lot of the things that I do, I can see the psychiatry note, right, I can see palliative care notes, I can see whether the patient's been dose reduced on their medications, I can see what side effects they're experiencing, like I can really get the whole patient. (BHP14)"*

One BHP has felt that the registry has improved their caseload review experience: *"So I like that we can be looking at the registry together, we can quickly at a glance, see what's a patient's oncologic situation. We have a list of medications they've been on, we can review pH nine scores, GED, seven scores, mood logs, activities together, and then we can really readily access. Okay, well, what* *did you tell them in your last session? What did I recommend at the last case review? So I think that is actually a more effective way to collaborate. Thus far, I think I am enjoying it with social work teaming, enjoying it. (BHP16)"* It also helped them manage their time better: *"So one thing I love is the initial scope registry face page where I can see all the patients enrolled, I can filter them by social work, I can get a quick snapshot of how severe their scores are. And really importantly, when the last time I reviewed their case for social work is okay, because we're still early in enrollment, I'm reviewing people too often, if anything, like more often than we need to strictly, then that will change. But it's really helpful to see that. (BHP16)"*

The registry introduces additional work, but there is a silver lining.

> *"My genuine belief is that with the learning curve, overcome, this is a better way to track patient data. But it's a real change from the way they've been doing things. And more than a change, it ends up being double the work because it's two separate systems that they have to nav compared to the one that they had to do.(BHP16)"*

Some BHPs secured a dedicated time to do administrative work *"that's specifically supposed to be set aside, to follow up on [their] registry and keep track of [their] patients. (BHP8)"* Still, many of the BHPs expressed that the system introduces more work to their already overwhelmed work. One major source of the frustration was the duplicated data entry in the registry and the EHR: *"I think one of the challenges with the registry overall is that, that it's disconnected from the medical record. And so we can't kind of pull in stuff that already exists in the medical record. I mean, just quickly toggle between the two necessarily, but I think that's just kind of the nature of the beast at this stage in the game. (BHP5)"*

Although integration with the EHR was desired, they mainly wanted data to be pulled from the EHR into the registry and not the other way around. One benefit of having a separate system is the ability to keep clinician notes separately from what patients have access to: *"In [EHR], we're sharing all of our notes with our patients. And so I'm, I'm more inclined to write less there and write more in this. (BHP1)"* BHPs mentioned that the information that they would put in the registry *"might not be something that [they] necessarily want the patient to read (BHP5)"* and need *"to be really careful now about what [they] put in [EHR] notes. (BHP4)"* BHPs were concerned with the EHR notes might be used against them and could be *"demanded in court. (BHP2)"*

Review: facilitating data-driven conversations

BHPs want to integrate the data into their treatment.

> *"Having that continuum of the data and the numbers to quantify things has helped to figure out what's working and what's not for the patient to be able to problem solve. (BHP15)"*

BHPs liked the assessment and mood visualizations in the registry and wanted to share the visualization with patients. Without overwhelming the patient, one BHP wanted a separate space that patients can go to for charting their data: *"So they have the option, but it's not necessarily in their face. So I wonder if there's even like a full another space here where you could do graphs. That's optional. (BHP11)"* Several BHPs wanted to view the visualization together during sessions to now only show progress. One BHP also wanted to celebrate small wins and validate their feelings: *"I like this idea, too, if it would be possible to share with them like, wow, I know, you're still really feeling like you're not feeling yourself, but look at where you've come from, it kind of helps using that to help them reflect. (BHP7)"* Several wanted to use the data as evidence to challenge patients' beliefs and perceptions about themselves: *"If they're telling me something, and the numbers say something different, right? They may say, I don't feel very, like I'm making any progress. And then you can show that will actually your scores are showing that there's even a small amount of improvement. (BHP11)"* A different BHP wanted to show correlations between symptoms and external events to help them prepare for the next such events: *"So I can say, Oh, look, that happened. It was you know, like, maybe you had a scan. Yeah. And they can see Oh, okay. Before scan business, I need to start doing some of the interventions and stuff like that. (BHP9)"*

Having granularity of data about their care practices can help BHPs problem-solve and modify their own treatment. For example, one BHP noted that knowing the frequency of use for different behavioral activation components would help notice that there is an issue: *"I think it's worth knowing how many times you go over stuff, because then you can notice if there's some kind of cognitive issues, or if for whatever reason, whoever you're working with is just not engaging in it. (BHP4)"* If they can notice the issue, BHPs can brainstorm solutions: *"this would be something that we've used? Well, we meet to look at what worked well, what didn't work? Well. What kind of adjustments we need to make? You know, are there barriers getting in the way? Do they need some motivational interviewing or something else? (BHP8)"* The treatment data can also be correlated with the outcome for the consulting psychiatrists to make treatment adjustment recommendations: *"I like the graphical representations. I like ability to see concurrent GAD-7 and PHQ-9 score. So like on a certain date, what's happening? I like the ability to be able to if I want to correlate those dates with Okay, well, what changed? Did we recommend that increase? Are you doing more behavioral activation? So, and I like the granularity. (BHP16)"*

Providers needed additional support for visual cues and data interaction capabilities to draw additional insights from the data and for interpreting the incoming stream of data. For example, they wanted various sorting and filtering capabilities: *"If there would be any advantage to sorting items by pleasure and accomplishment. Okay, so we can see that these particular activities that you did seem to be high pleasure and or low pleasure activities. Again, you know, either sorting or color coding or something so that there could be kind of a visual cue of some kind. In the event that either I'm looking at this on my own trying to suss out that pattern, we're looking at it with a patient trying to show them a pattern. (BHP12)"* BHPs felt that it is a learning curve to figure out how to make data actionable: *"It's just going to be a clinician learning curve for me to figure out. Well, on October 4, they recorded you know, this thing on the GAD and this other thing on the PHQ-9, and now I have mood ratings on that same day. So now I've got to kind of figure out, what does all of this mean, right? It's going to be a learning curve. [...] What do I do with information that comes in? (BHP9)"*

Patients want guidance on what to do based on their data.

> *"If I'm between moderate and severe [depression] that, you know, what can I do to improve? (Pt9)"*

Many patients were unsure of what actions to take when they knew that their assessment scores reflected poor mental health. Patients desired the app to give them guidance on how to change their mood from bad to good (Pt4), to prompt them to brainstorm ways to improve their score (Pt9, Pt7), or to provide inspirational quotes to help motivate them (Pt7). One patient made a recommendation: *"In more crisis or a more acute mode, [...] add clinicians recommendation. (Pt8)"* Another patient recommended that an inspirational quote following a bad assessment could include an action:

> It's quote, actions. So he gives you a quote from someone like this one [...], 'My therapist told me, the way to achieve true inner peace is to finish what I start. So far, I finished two bags of m&ms and a chocolate cake. I feel better already'. Now, this was a quote [...] from a comedian. And then [...] the action part of that is, your action for today is to treat yourself to a small indulgence, you know, [...] there's a quote, and then there's an action. (Pt9)

Plan: documenting care plan during and outside of sessions

Patients need guidance for using the app.

> *"Whether we can get them to use it correctly is another thing, but I'm happy to work on that. (BHP7)"*

Patients sought guidance from their providers on how to incorporate the app into their care plan: *"You need to walk me through this stuff for me to kind of get my brain wrapped around. (Pt14)"* Values and activities inventory stood out as one of the activities that patients would need guidance on the most. One BHP explained that her standard practice for the paper version of values and activities inventory is to start the exercise together: *"I'll often do like one part of it with them just to kind of give them you know, an idea of, of, of, you know, its purpose and how to do it, and then they'll usually take it home and fill out the rest of it. (BHP12)"* One BHP explained that, in her previous experience, patients *"sometimes mistake activities for values and vice versa. (BHP16)"* Instead, a digitized version of the values and activities inventory, while readily available, lacks *"enough direction for them to understand what [they]'re asking. (BHP16)"*

Guidance on how to use the app or how the app integrates into the care became especially salient during the later stages of our user study. All four providers in the pilot study expressed the need to first walk patients through the app. However, this became a logistical challenge, as one BHP explained: *"We keep trying to meet face to face so I can kind of show her. And yeah, like, I can't explain it on the phone. And she's 70 something and she's like we just can't work out a time when because I'm only in the [redacted] clinic one day. Yeah, that never seems to work for her. (BHP13)"* In addition, the BHPs need to understand the app to be able to guide patients. One BHP said, *"I also feel that I don't know the app inside and out. So I, okay, I stumble a bit in trying to teach them. So even the times that we tried to do it together, you know, I've pulled out the paper instructions of how to do that. I've given them a copy of it. But I don't know. I don't know that that's very helpful. That might be just be too much. (BHP15)"* Another BHP explained that the app needed to be gradually integrated into the treatment work flow: *"What's going to slow down my ability to do BA is that I'm going to have to incorporate this app and start slow. (BHP9)"*

Integrating the system into the care introduces a challenge for the BHPs. Typically, BHPs do not bring technology into their session. With the system, BHPs envisioned that they would have to change their approach to sessions: *"it sounds like the idea is, maybe that I would be I would be on my my computer during the session with the patient. Yeah. Yeah. And that that will be a new, a new approach. (BHP12)"* One BHP was *"willing to change workflow if it helps patients [and] can bring about positive change for a patient. (BHP9)"*

Having activities documented can help patients be accountable.

> *"Keeps one accountable here, doesn't it? (Pt11)"*

On the other hand, patients in general had positive feedback about the activity scheduling and logging experience. They appreciated the flexible support around activity scheduling and logging which increased their appeal for the task. One participant especially liked the ability to think proactively about alternative activities in case the situation does not allow for the target activity: *"Instead of saying, Well, I just can't do this. Having, you know, kind of put it in front of me to Okay. Think of something you can commit to that would you know, fill in the blank? (Pt5)"* Tracking alternative activities would help *"pinpoint it a little bit more (Pt5)"* for what is really getting in the way.

Several patients desired the scheduled activities to be directly integrated into their calendars, citing that *"If it's not on my calendar it doesn't exist. (Pt14)"* Because patients often had multiple medical appointments, they wanted the activities to be merged with the appointment calendar maintained by the EHRs (Pt6).

However, one participant emphasized that flexibility in the timing of activities and reminders are important in getting them done in case of unforeseen disruptions that happen frequently to patients with cancer: *"Couple of things I would say as someone with a lot of medical things going on, maybe having an activity that is associated with a specific time is not going to be very practical. And if I'm thinking oh, I have to do this at this time, and if I don't do it at that time, I might not do it at a different time. Yeah. And you know having more latitude or variability just like okay, and I do understand when it is routine, there is huge value in that in the routine it becomes more of the fabric of the course of your day. But when that routine gets upset, for whatever reason interrupted I would hate for the activity then to fall by the wayside. (Pt8)"*

Patients thought that reminders also helped nudge them to do important activities when there is no social support: *"When I first got treatment when I have a friend who* *come and stay with me for a week, so even my husband go to work and my daughter go to school, there's somebody reminding me to eat it reminding me to work out reminding me for a walk, it makes a difference. So this is like you don't have somebody that you every day. It's like important, especially for cancer. (Pt4)"*

One type of appeal for activities being tracked in the app is the potential for the reminders to be personalized, especially by their BHPs, to ease the sense of being overwhelmed by their situation: *"If your app could send you a little notifications, like, Hey, [SW] says, Don't forget to do your check in for the day. (Pt15)"* Patients envisioned that different people would want different modes of reminders (e.g., text messaging, in-app notification), different messaging in the reminders (e.g., personalized, anonymous, coded), and different timing of reminders (e.g., random, specified) to maximize the uptake of activities: *"I think I feel like people will probably want to different people would want different things. [...] If it showed up at two o'clock every day, I would get used to it. And I think it would be way too easy to ignore it. (Pt5)"*

Because the goal of the app is to increase patient activities, some wanted nudges during a period of inactivity: *"if you haven't done a lot of stuff. Along those same lines, like how are you doing? Are you sure there's no, like activities you want to log? Like, if not, here's, here's something you could do. Maybe make when people at some point near the beginning, make people generate ideas for what they can do? (Pt2)"*

Closing the loop: enabling collaboration between sessions

BHPs can interact with patients in real-time.

> *"I think one thing that it seems like will be important is for people to feel like I'm actually paying attention to the stuff that they're doing. [...] I think that that's going to be one of the key features of this for us is being able to act more in real time. (BHP12)"*

Because our system does not have an active alerting system for providers, BHPs were concerned that they would not be viewing the registry actively to react to any incoming data that might raise a flag (BHP7). Without a sophisticated alerting system, one BHP suggested sending an email: *"It wouldn't even have to include information about the patient's name, just like, Hey, your registry is on fire right now. Go check your registry. (BHP7)"* Many BHPs wanted to use the data to reach out to patients who might be having a hard time: *"I can see if there's really drastic things and I was had safety concerns. I could see kind of ad hoc adding an appointment or just a casual phone call or something like that to check in. (BHP7)"* If a patient was doing better and engaging in activities, one BHP wanted to reach out with a *"small cheerleading note"* or a *"note of encouragement. (BHP12)"*

One BHP thought that the system would help them make faster progress compared to traditional one-on-one sessions: *"That's another big limitation to our work is that we see patients in the clinic, you know, half an hour an hour every couple of weeks. And I can say multiple my current patients, I if I had a way of reminding them or giving them assignments, we will be making much faster progress or for some people any progress at all because they leave and they're not talking to me and their mind is just you know, moved on, which I would say for myself to do. So this is great. Like this. (BHP7)"*

Patients want flexibility in sharing their data with providers.

> *"One of the few things that like we as patients have control over a lot of the time, is like what we tell our health care providers. (Pt2)"*

Most patients had differing opinions about sharing their data with their providers. Some were sure that the emotional data tracked by the app would not be useful for the oncologists (Pt8, Pt14). Some were open to sharing their data with the care team because they *"feel so bad that [their] privacy was not an issue (Pt6)"* or because the information would be *"relevant to [their] condition (Pt8)"* and *"so that they can help. (Pt6)"* One patient wanted the data sharing to be flexible so that they can choose *"one person kind of as like liaison [...] to look at my general, like all my info reported. (Pt2)"* Some wanted to share data on an agreed upon schedule, such as once a month (Pt11) or every few weeks (Pt12).

Some were hesitant about sharing the data right away or being honest about their data. One patient said that they may not submit the data: *"If it gets that bad, I start to go into shutdown mode, and I don't want to talk to anybody might not even hit that button. (Pt9)"* Another patient was fearful of sharing their perceived failures: *"It was like when I had to do my socialize every day. I mean, it was so hard. I didn't want to tell her I didn't do it three days in a row. You know, I mean, ya know, in some days when I come when I did, I... I guess I felt better, but it was scary. Yeah. You know, um, so it's just, I... It's hard to say, you know, I mean, if, if you face any challenges, I think it's just that commitment to fill it out correctly. (Pt3)"*

Patients want assurance that there is someone on the other side.

> *"It would be really good if this was a way in which, you know, maybe the PA can say, we got your back. [...] or [SW] will say, [...] Hey, we're just checking in on you, and you're back. And, and is there anything you need? (Pt10)"*

Having the tracked data sent to the providers meant that patients had an expectation for responses or acknowledgment. In extreme cases, some patients wanted a response *"as soon as [the data is] available (Pt12)"*, but most were comfortable having providers view the data on a regular cadence unless there was an emergency. The expected cadence varied from daily (Pt6), a few times a week (Pt9), to weekly (Pt6). One participant expected that it would vary depending on the severity of the symptoms: *"If I was in some kind of an acute or crisis mode, I would assume they're going to be* *checking more frequently. If a fire if things are cruising along, maybe they're not going to need to access that data. (Pt8)"* Some patients were fine reaching out to providers if there is an urgent need (Pt8, Pt11). At the very least, patients expected the providers to review the data before their appointment: *"My hope would be that they would review that information before the appointment Tuesday or Wednesday before I come in so that we're, we're current. (Pt7)"*

Many patients wanted the BHPs to reach out immediately if the assessment scores were bad, if suicidality was detected, or if a consecutive log of a bad mood or severe symptoms. Patients expected BHPs to notice that something is not going well: *"If i'm going down hill in the rapid way ,it'd be really good for <SW> to know that no and not me having to kind of reach out on you know. (Pt10)"* If the depression score is severe, some patients wanted the system to *"alert the care team right away (Pt12)"* with a *"red flag (Pt6)"* and for the BHPs to respond *"aggressively (Pt14)"*: *"Definitely, definitely. You know, them to reach out as soon as possible to see how I'm doing? I'm not the type of person that would go and hurt myself. But there are times where I just feel like okay, especially right after the surgery, [...] there's just too much too much going on.(Pt9)"*

One patient wanted a way for the system to let them know what the committed response time would be: *"I think as a user, I would want the other options to be on the provider side to say, Okay, I'm having my appointment, and we discuss and we decide that he, for me, he's only going to check this my data right before my appointment. So he has some way to designate that on my record, whereas his next patient comes and maybe they're in a tougher place, and that maybe he needs to be able to see that on an ongoing basis. And that he has some way to designate that on their record. (Pt5)"*

4.2.5 Discussion

In this study, we designed and evaluated technology-enhanced collaborative care and behavioral activation system that improves stakeholder collaboration and engagement in the care of patients with cancer and depression. Our design was guided by tasks identified from our model for technology-enabled collaborative behavioral activation intervention loop (Figure 4.5) and iterated through 3 user studies and a pilot deployment study at multiple cancer clinics. Analyzing the interviews from these four studies, we found that our system supports all four touch points identified in our model with the potential to bring positive changes to the existing care by closing the intervention loop.

We found that patients saw the app as a way to engage in care between sessions by committing to mood-enhancing activities and tracking progress toward them. Patients reported an increased sense of accountability and self-efficacy in lifting themselves up between sessions. For BHPs, the registry provided a clear structure around their treatment in reminding them of various treatment components that they have not exercised or organizing information for efficient reviewing of the patient progress. It also provided a safe space for BHPs to document their notes that may be relevant to their treatment. Both patients and BHPs felt that the system would allow collaborative review of the tracked data and problem-solving to brainstorm the next steps during the sessions. Outside of the sessions, BHPs thought that the system would allow real-time monitoring of patient's progress and engagement.

Our study revealed several opportunities for improving the system as well as serious challenges in integrating the system into care practices. Although some challenges could be attributed to the system design, many challenges were brought on by the introduction of the system. Such challenges must be resolved through careful deliberation by the stakeholders about the role of the system and their motivation to integrate the system into the care flow and adapt their approach to care. Here, we discuss both technical and non-technical challenges and opportunities for future design and research.

Support personalization of the patient care experience

Many patients expressed the desire to personalize the app. In addition to the personalized touch to the background or the color scheme, patients wanted the app to help them understand themselves better. Patients wanted the ability to adapt the app to track their mood, activities, and other symptoms the way they see fit. Patients also expressed that their tracking needs to be varied based on the cancer stage, the treatment they were in, or their underlying depression. Therefore, a future collaborative behavioral activation system should provide flexibility in tracking. For example, the system could allow customization of the tracking forms based on how patients want to articulate their mood or activities (e.g., numeric rating, emotion words, emojis, or diary entries) as well as other relevant data (e.g., sleep, pain, nausea) that could inform their mood or activities entries. Such customization for tracking should be discussed with the providers so that the data collected is clinically relevant [362].

Both patients and BHPs in our study expressed the need to draw insights from their data, but the kinds of insights they wanted were diverse and depended on the kind of questions that the care team needs to answer (e.g., What kind of activities lead to positive mood? What time of day do I feel low? What gets in the way of completing certain activities? Do I really feel miserable all the time?). Simply presenting visualizations or insights dashboards is insufficient if the questions or tracking goals are not documented, especially for patients with cancer and depression who may suffer from cognitive or memory issues. Therefore, a collaborative behavioral activation system should allow patients and providers to set specific questions or tracking goals that are relevant to the current treatment. The system could also allow the design of custom insight generation tools (e.g., visualizations, summaries) that would help patients and providers to directly answer those questions [361].

Patients often had great insights about how technologies should be customized to better integrate them into their lives, especially around nudges. We found that different participants had their own ideas for acceptable notification designs that would support their activities and tracking based on their current use of technology. For example, some patients wanted reminders to come as SMS messages while others wanted notifications at random times throughout the day or wanted them to be in an actionable language. Although there are standard conventions for notifications and reminders on various mobile platforms and calendaring systems, it is important to remember that the nudges to track or perform activities in our context come from a treatment plan that patients and providers agree on. Therefore, a collaborative behavioral activation system should provide flexibility in designing notifications and reminders but also such design decisions should ultimately stem from shared decision-making between patients and providers.

Balance self-care and collaborative care

The primary design objective of our collaborative behavioral activation system was to improve stakeholder engagement and patient outcomes through technology-facilitated collaboration. This is why our system does not necessarily provide stand-alone, self-help activities to patients nor does it provide just the charting functionality in the registry. As we mentioned earlier, the system integration fails if both patients and providers do not engage in the process, but should the engagement level be expected to be consistent? Our study revealed that there is a variable level of engagement needed by patients and providers depending on patients' prior experience, self-efficacy, and where they are in their treatment. We saw evidence that patients loved the ability to engage in self-care between sessions, but they wanted BHPs' guidance through that process at the beginning. Patients that have prior experience in mood logging or journaling expressed confidence in collecting data on their own and analyzing them to draw actionable insights, while those without the prior experience needed encouragement and direction. When patients are in challenging situations (e.g., during chemotherapy, receiving bad news after a scan), they wanted more proactive engagement from their BHPs.

Our current app design includes behavioral activation components that can stand on their own and be utilized by patients at any time. But having content available prematurely without clinical guidance can lead to patient confusion. For example, the values and activities inventory, which is one of the earlier components in the behavioral activation treatment and one that we digitized and made available in our app, was seen as one that both social workers and patients preferred to start during the session. Providers expressed concerns that patients would not complete the exercise properly or comprehend what is being asked of them. We also know that in-person sessions may be too few and far between for patients with cancer and depression due to the various challenges they face with their care. So inevitably, patients may need to rely more on self-care activities and self-help content without direct guidance while BHPs monitor and check in from a distance. Not having access to patients or patients conducting behavioral activation activities on their own poses a risk of low adherence to guidance-level care. Therefore, further research is necessary to understand how to design for flexibility in the level of engagement by patients and providers but also the impact of flexibility on the quality of care and adherence.

Support negotiating the right level of data sharing and responsiveness

In our study, we found that patient expectations for provider responsiveness towards the incoming stream of patient-generated data ranged from immediate to the next session. We also found that patient comfort levels in sharing data varied from full transparency to full control. Patients raised concerns about knowing if the data was received, and providers raised concerns about liability if they were not able to keep up with the incoming stream of data [109]. Beyond the transfer of patient-generated data, digitization of care could enable a certain level of visibility [236] into the care practices (e.g., when BHPs logged in) as well as the patients' engagement (e.g., when patients logged in) that clashes with a level of autonomy, control, and confidentiality desired by patients or providers [53]. Such tension among control, visibility, and accountability, if not balanced, could erode the trust in patient-provider relationships [376].

Our findings highlight the need to establish a process for patients and providers to negotiate an appropriate level of data sharing and responsiveness before introducing the technology into the patient-provider collaborative relationship and a process for making adjustments as appropriate [454]. Such shared decisions must be documented and respected by the collaborative system. However, such processes could be in conflict with the organization's need for standards that must be put in place to meet the legal and regulatory requirements surrounding data ownership, privacy, and security [231, 313]. Future work is necessary to understand the individual, interpersonal, and organizational needs for setting appropriate boundaries for the collection of patient-generated data and how to design processes around them.

Compensate for extra work through increased efficiency and productivity

Even though the system has the potential to enhance the sessions and to fill the engagement gap between sessions, our findings revealed that there was a significant cost incurred by the introduction of the system. Many patients were concerned that the app felt like more work on top of everything that they are already dealing with. Despite the additional work, patients were generally willing to engage if there were sufficient justifications from their BHPs. Some patients could not imagine engaging with the app when they are in "survival mode" or hospitalized. Social workers also did not have high expectations for app engagement for patients whose health was really bad. This observation raises the question of the target audience for the system. Behavioral activation may be the most effective treatment for depression in patients with severe health problems, but perhaps getting these patients to engage in a technology-facilitated behavioral activation may not be as beneficial because of the additional cost incurred. Further research is necessary in evaluating the cost-benefit trade-off of introducing collaborative behavioral activation systems for different patient categories, for example, along dimensions of depression severity, cancer stages, or financial or logistical hardships.

The costs to BHPs were felt from multiple perspectives. There were administrative costs associated with duplicated data entry from having a separate system from the EHR and with monitoring the incoming stream of data from patients for any urgent and necessary follow-up. Costs associated with changing their care practices were also high. BHPs felt that they had to incorporate behavioral activation even when the more flexible therapeutic approach was deemed necessary, they had to bring technology into the session when they were used to more traditional face-to-face interactions, and they had to be able to interpret and utilize new forms of patient-generated data beyond standard scales [232, 109]. There may be technology-based solutions to reduce the cost by gaining efficiency through the registry. For example, patients who need urgent attention could be flagged for proactive outreach before situations escalate into a crisis. Integrating the registry with EHR and adding the ability to keep private notes, as many social workers desired, could eliminate the cost incurred by duplicated entries. Even though the system might improve patient outcomes overall, efficiency gain in one set of tasks (e.g., patient flagging) may not be enough to offset the cost incurred by the introduction of new tasks (e.g., monitoring patient-generated data, adapting the care practices to integrate technology) for an already overburdened BHP population. Therefore, the integration of a collaborative behavioral activation system should be evaluated on patient outcome metrics (e.g., engagement, depression) as well as metrics specific to BHPs that evaluate their productivity (e.g., patient crises averted, cases reviewed) and well-being (e.g., work satisfaction, perceived workload, provider stress or burnout).

Resolve social and organizational tensions surrounding technology integration

We identified several challenges that prevent getting stakeholders started in integrating the system into their care. The first challenge is getting set up with the system. BHPs were not equipped to explain how to use the app or troubleshoot the app with their patients when the patients were getting set up with the app. They were not able to walk through the app over the phone when in-person appointments were not possible. This challenge could be addressed by designing a walk-up-and-use system or incorporating self-guided tutorials. However, this places the burden on the patients who are already overwhelmed with their cancer and are impacted by the side effects of the cancer treatments.

The second challenge is in getting patients to be motivated to use the app. Patients wanted to be convinced of the value of using the system by their BHPs and wanted to see evidence that the data that they collect was actually being used to inform care. In fact, goal alignment and shared understanding of how the data will be informing care is critical in patient-provider collaborations with patient-generated data [89]. Ultimately, the burden of introducing the app and getting patients on-boarded falls on the shoulders of the BHPs who are already overburdened with other tasks. However, some BHPs were hesitant about the use of the system in their care, citing patients' disinterest and complex situations or the time spent on introducing the system away from therapy.

This creates an impasse. If BHPs do not introduce the app to the patients, patients do not engage in the app, and BHPs, as a consequence, do not engage in the system. Even if patients engaged in the app, if BHPs do not incorporate the data in their care, patients will eventually abandon the app. Both patients and BHPs are equally overburdened and under-resourced to adopt new care workflows. So, how do we address this problem? Is it the problem of the system designer to make the system persuasive? Is it the patients that should figure out the app on their own and demand its use in care? Should the BHPs spend their precious session minutes selling the benefits of the app?

Although our system was recognized as potentially beneficial to the care, our findings revealed that there may be social and organizational barriers to adoption that may not be solved with technology innovation alone [232]. Future research should, therefore, understand what changes to the organizational culture, care delivery, and team composition need to be established when a new technology is introduced. For example, an in-house technical expert can troubleshoot technology issues with patients and providers. Standardized training and guidance can help BHPs identify patients that might benefit the most from technology integration, help with getting the initial buy-in from the patients, and establish a playbook of use case scenarios of when and how to incorporate patient-generated data into the care. There may need to be a drastic shift towards the organizational culture that values the use of patient-generated data as an integral component of delivering high-quality care [454].

Investigate facilitators and barriers for technology adoption

Our collaborative behavioral activation system is currently being evaluated in a randomized clinical trial with two arms, where patients are randomly assigned to technology-enhanced collaborative care (i.e., where our system is used) or usual collaborative care. The RCT as a whole would evaluate whether technology-enhanced care improves patient engagement and outcomes across the population. The randomization of patients brings an interesting challenge to the adoption of technology because not all patients in the technology arm would be an appropriate candidate for the use of the behavioral activation app, just as our pilot social workers have expressed. However, it introduces an opportunity to identify patient characteristics that facilitate or impede engagement in technology-supported care and to evaluate the level of engagement necessary to observe clinical efficacy.

Our study also found barriers to the adoption of the collaborative behavioral activation system by both patients and providers. In evaluating the impact of these barriers to the adoption of the system, we recommend the Unified Theory of Acceptance and Use of Technology (UTAUT) [426] to be the most appropriate in capturing the needs of our population. For example, patients expressed that they would be motivated to use the system if their social workers justified its use, and social workers pointed to the organizational support (e.g., administrative time) as a facilitator for engaging with the registry. Both of these examples could be explained by the Social Influence construct. Facilitating Conditions is another construct that could explain the patients' hesitation to use the system when they are overwhelmed by multiple demands or when they are severely depressed. Therefore, we recommend measuring the core constructs of UTAUT at each iteration of the system or the implementation to understand factors that influence technology adoption and the efficacy of technology-enhanced collaborative care.

The integration of psychosocial services in cancer care is crucial in supporting the whole patient, and yet challenges rooted in cancer-centric views continue to prevent patients with cancer and depression from receiving adequate care. We introduce the parallel journeys framework to characterize patient experiences in simultaneously navigating the cancer care journey and the psychosocial care journey. The framework helps reveal challenges at the seams of the two journeys and helps surface related technology opportunities. We encourage health technology researchers, designers, and healthcare providers to consider both cancer care journeys and psychosocial care journeys in their future design of technologies for patients with cancer and depression.

4.3 SUMMARY OF CONTRIBUTIONS TO THESIS

The integration of psychosocial services in cancer care, such as collaborative care, is crucial in supporting the whole patient, and yet challenges rooted in cancer-centric views continue to prevent patients with co-morbid cancer and depression from receiving adequate depression care. In this chapter, we aimed to understand the challenges of depression care in cancer settings to inform the design, development, and deployment of technology solutions for supporting collaborative care.

We conducted a formative study to find that patients with co-morbid cancer and depression struggle to navigate between their cancer and psychosocial care journeys and proposed the parallel journeys framework as a conceptual design framework for characterizing challenges that patients and their care teams encounter when cancer and psychosocial care journeys interact. We used the challenges discovered through the lens of this framework to highlight and prioritize technology design opportunities. By digitizing the core components of behavioral activation and collaborative care, we developed a model for a technology-enhanced collaborative behavioral activation loop. Using this model, we designed a technology-enhanced collaborative behavioral activation system, called SCOPE, and evaluated the system through three design iteration studies and a pilot deployment.

We found that patients generally welcomed the technology support but needed a clear justification and a collaborative commitment from their BHPs to adopt its use. The system enabled patients to engage in self-care activities between sessions and helped BHPs by providing structure around their treatment and caseload reviews, with the potential to facilitate data-driven conversations during the sessions for collaborative action planning. However, we found that the introduction of the system added friction in existing clinical workflows and burdens that got in the way of both patients and providers engaging in the system.

The parallel journeys framework introduced in this chapter helped raise awareness that mental health care is deprioritized over cancer care and that tools and processes need to be implemented to improve engagement in mental health care. However, our findings also revealed that introducing technology enhancements to collaborative care, although informed by the needs discovered through the lens of this framework, still faces technical and non-technical issues when it meets the complexities of the clinical real world. Our findings highlighted the need to overcome several implementation challenges such as inadequate negotiation between patients and BHPs before the technology introduction, lack of knowledge about how to incorporate the system into the existing care practices, and motivational and situational barriers to technology adoption. Therefore, this work suggests that efforts to introduce technologies aimed at improving collaboration and engagement among multiple stakeholders in clinical settings must be prepared to make appropriate adjustments to the interventions, the process, and the organizational structures surrounding them.

My work in the clinical context of cancer and depression demonstrated that *human-centered methods* – formative, design, and deployment studies – can reveal *multistakeholder and organizational challenges* that arise at the intersection of cancer and psychosocial care journeys as well as when new technology systems are introduced to the clinical care context. To *improve the collaboration and engagement in depression care*, I *digitized core components of behavioral activation and collaborative care* into a collaborative behavioral activation system. From the deployment of the system in a real clinical setting, my work further revealed opportunities for *adaptation in both technology enhancements and implementation processes.*

Chapter 5

INTEGRATING STRESS-REDUCTION INTERVENTIONS IN THE WORKPLACE

In the previous two chapters, I demonstrated that using human-centered and computational methods can reveal human needs and challenges in the context of the pandemic and in the clinical context of depression care. I also demonstrated that digitization and adaptation of core components of evidence-based care strategies can help improve engagement and collaboration in depression care for patients with co-morbid cancer and depression. So far, I have leveraged the contextual information at extended temporal scales (i.e., years, treatment duration). In this chapter, I situate my research in the context of workplace stress to explore how *momentary* contextual information can be leveraged to adapt interventions and to improve engagement.

Stress is prevalent and costly: Over 83% of Americans suffer from work-related stress [293] and over half experiencing stress throughout most of the day [328]. Workplace stress intervention strategies have been recommended and evaluated [98], but integrating individual-based stress management interventions into the workday and actively engaging with them can be challenging [123, 184]. Smartphone-based mental health apps have gained popularity for ease of access, and recent efforts have been focused on improving usability, delivery timing, and content selection factors to improve user engagement. Bite-sized, "digital micro-interventions" [36] have been a promising content design approach and can be easily integrated into the workday. To further improve engagement, just-in-time adaptive interventions (JITAI) promise to deliver the right personalized micro-interventions at the right times to maximize efficacy [292, 355].

In this chapter, I focus on workplace stress as a context for which to design stress-reducing micro-interventions. I demonstrate that passive-sensing and momentary contextual information can be leveraged to achieve *just-in-time adaptation of interventions* aimed at *improving engagement in stress-reduction interventions* in the workplace. I conduct two studies to illustrate that deploying an adaptive workplace stress intervention system can lead to workplace stress reduction and help identify factors of engagement.

The first study [190] examines the impact of the intervention types and delivery timing through direct integration of stress-reduction interventions into the work context (Section 5.1). We designed and developed a workplace stress intervention system that uses passively sensed contextual data to deliver just-in-time digital micro-interventions, adapted from components of Cognitive Behavioral Therapy (CBT) and Dialectical Behavioral Therapy (DBT). Using this system, we conducted a four-week longitudinal study with 86 participants, examining the effects on usage, stress reduction impact, and user preferences by three intervention types and two delivery timing conditions: *Pre-scheduled* (PS) by users and *Just-in-time* (JIT) prompted by the system-identified user stress-levels (Section 5.1.1). Although we found no significant difference between JIT and PS conditions in study-long or momentary stress reduction, participants preferred automated "nudges" over scheduling their own interventions, while simultaneously desiring control over their schedule with system assistance for intelligent planning (Section 5.1.2). Our findings suggest that users may benefit from a combination of the two delivery methods, wherein system-initiated interventions offer ease and increase overall usage, and user-initiated, pre-scheduled interventions promote a sense of control and could lead to healthy behavior change. This study provides opportunities to guide the design of personalized JITAI and planned intervention systems to reduce stress and enhance well-being in the workplace (Section 5.1.3).

Using the data collected from the deployment study, the second study examines what factors contribute to improving intervention engagement and positive outcomes to inform the design of JITAI systems (Section 5.2). We analyzed the interaction data from JIT participants together with the contextual information captured by the sensing software and telemetry as well as demographics and individual characteristics (Section 5.2.1, Section 5.2.2). We built statistical models to identify contextual, individual, and intervention-related factors that explain the likelihood of choosing different intervention types, engaging in a system-initiated intervention, rating the intervention, and the effectiveness of intervention in momentary stress reduction (Section 5.2.3). We found that personality traits influence not only the engagement in system-initiated interventions but also in the choice of interventions, suggesting that a JITAI system may need to adjust its recommendations based on personality traits to balance choice preferences but also engagement likelihood. We also found that contextual factors, such as the timing, being in a meeting, the availability at the computer, or the appropriateness of the intervention in the given context, were all important to consider for engaging in system-initiated interventions, but the effects of intentionally deferring the intervention to a later, specified time were much higher. The study findings suggest design implications for tailoring JITAI systems to meet the contextual and personalized engagement needs that balance user preferences and intervention efficacy (Section 5.2.4).

The key takeaways and contributions to my thesis from these two studies are summarized in Section 5.3. The first study was conducted in collaboration with Marah Ihab Abdin, Mary Czerwinski, Javier Hernandez, Esther Howe, Daniel McDuff, Mehrab Bin Morshed, Gonzalo Ramos, Kael Rowan, and Tracy Tran and was published in CHI 2022 [190]. The second study was conducted in collaboration with Tim Althoff, Mary Czerwinski, Javier Hernandez, Esther Howe, Robert Lewis, and Koustuv Saha and is yet to be published. I contributed to the conceptualization of both studies, the design and development of the system, the data collection and analysis, and the interpretation of results. This work is currently in submission.

5.1 PROVIDING INTEGRATED SUPPORT FOR STRESS REDUCTION AT WORK

Reducing workplace stress is of critical importance for employees and employers alike. Work-related stress increases the risk of mental and physical health disorders, decreases productivity due to absenteeism and burnout, decreases overall job satisfaction, and increases rates of stress-related accidents and employee medical, legal, and insurance costs [55, 444, 62, 94]. Workplace stress can also spill over into life outside of work, disrupting the overall well-being of workers [162]. Workplace stress intervention strategies such as organizational changes, individual stress management skills training, and therapeutic counseling have been recommended and evaluated as important components of long-term stress reduction [98]. Individual-based stress management interventions (e.g., cognitive-behavioral skills training, meditation, exercise, etc.) have been shown effective on psychological, physiological, and organizational outcome measures [333].

However, integrating individual-based stress management interventions into the workday and actively engaging with them can be challenging in workplace cultures with psychologically unsafe climates [123] or where there are high task demands but taking personal time for stress management is not supported [184]. Furthermore, while employees may learn about stress management strategies via in-person stress-reduction workshops or individual therapy, shifting from *understanding* evidence-based strategies to *using* these strategies when they are most needed (i.e., in moments of high stress) can be quite difficult psychologically, even in healthy workplace contexts that value employee well-being [179]. Smartphone-based mental health apps have gained popularity because they are always available (unlike a therapist), and can be used without drawing attention to the user or disrupting a work environment. However, most available mental health apps are not evidence-based [192, 294], and those that have been rigorously studied – although shown to be effective in improving relevant outcomes [225, 244] – also have high attrition and low adherence [243, 134]. Therefore, promoting active engagement in interventions over time is an important step to improving user engagement and long-term outcomes for technology-delivered interventions [292, 305].

Recent efforts to improve user engagement for technology-delivered interventions have been focused on improving usability, delivery timing, and content selection factors. Bite-sized, "digital micro-interventions" have been a promising content design approach to lowering the barrier to entry and to reducing the effort needed to engage with the content [36]. Because such interventions can be accessed via computers at employee workstations (i.e., do not require context switching), they may be easily integrated into the workday. To further improve engagement, adaptive or personalized micro-intervention delivery systems can incorporate individual preferences and contexts to choose the appropriate intervention content and delivery timing [218, 305, 365, 121]. For example, just-in-time adaptive interventions (JITAI), powered by ubiquitous sensing technologies, promise to deliver the right interventions at the right opportunistic or vulnerable times to minimize disruptions and optimize efficiency [292, 355]. Just as therapists select the most appropriate intervention for a given moment, users could benefit from the personalization promised by JITAI systems [292]. Although JITAI has the potential to become more intelligent over time and improve its predictive ability to identify opportune moments for intervening [355] and the ideal content, most conceptualizations of this approach do not include opportunities for individuals to *proactively* engage in and exercise control over their stress management processes, which has been shown to improve long-term psychological outcomes in individual psychotherapy [207, 93] and to lead to health behavior change [396]. Accordingly, our research examines the role of the system- and user-controlled intervention timing and content selection in promoting engagement and improving stress-reduction impact.

In this chapter, we present a four-week, between-subjects study with 86 information workers. Our aim was to understand the impact of digital micro-intervention *delivery timing* and *content* on usage patterns and stress reduction throughout the workday to inform the design of effective and engaging workplace stress reduction intervention systems. Leveraging a desktop application to facilitate passive data collection and a Teams chatbot for intervention delivery, our study compared three categories of intervention content and two delivery timing conditions: *Pre-scheduled* (PS) by users and *Just-in-time* (JIT) according to passively-sensed and user-reported stress-levels.

We found our interventions to significantly reduce momentary stress. Although we found no significant difference between JIT and PS conditions in study-long or momentary stress reduction, participants preferred automated "nudges" over scheduling their own interventions, while simultaneously desiring control over their schedule with system assistance for intelligent planning. While our users rated the shorter, "easier" interventions as more enjoyable, we found that the longer, more difficult-to-perform interventions were in fact significantly more effective. Our findings suggest that both system-initiated intervention delivery and user-initiated intervention scheduling are promising directions for integrating stress-reduction interventions at work. In fact, users may benefit from a combination of the two, wherein system-initiated interventions offer ease and increase overall usage, and user-initiated, pre-scheduled interventions promote a sense of control and could lead to healthy behavior change. In both cases, our findings suggest users benefit from having access to a healthy balance between easy-to-do and effective interventions. Based on these findings, we present opportunities to guide the design of personalized

JITAI and planned intervention systems to reduce stress and enhance well-being in the workplace.

5.1.1 Method

The goal of this work was to identify design opportunities for systems that integrate digital micro-interventions into everyday work contexts. Our research questions were:

RQ1. How does intervention timing impact intervention usage, stress reduction, and user preference?

RQ2. How do different types of interventions impact intervention usage, stress reduction, and user preferences?

RQ3. What aspects of the intervention timing and content do participants find most useful or needed?

To examine the impact of different delivery timings (RQ1), we conducted an experimental study that directly compared two conditions: *Pre-scheduled* (PS) by users and *Just-in-time* (JIT) according to user stress-levels. We designed and developed an intervention system with a chatbot that facilitated different delivery timing of stress reduction interventions to information workers (Section 5.1.1). In the PS condition, the chatbot helped participants browse through the catalog of interventions and schedule them in their calendars; in the JIT condition, the chatbot nudged participants to perform interventions when it detected high-stress levels. To examine the impact of different intervention types (RQ2), we adapted evidence-based psychosocial interventions into digital micro-interventions and categorized them into three types according to their function and required user effort (Section 5.1.1). Participants chose from the three intervention types throughout the study. Finally, to solicit user feedback from in-vivo usage of the system (RQ3), we deployed this intervention system to 86 information workers and conducted a longitudinal study with four weeks of observation that compared two delivery timings and three intervention content types (Section 5.1.1).

System Design

We designed and developed an intervention system that is composed of (1) a *stress score* component, (2) a *JIT* component, and (3) a *chatbot* component. Our *stress score* component computes the user's current level of stress based on passively sensed data. Then a *JIT* component leverages the stress scores and the user's self-reported stress levels to determine when to nudge users to perform stress-reduction interventions. Finally, a *chatbot* component proactively sends messages to the user and facilitates the delivery of ecological momentary assessments (EMAs), surveys, and intervention content. We designed our system to capture salient signals in people's work context where they are more likely to be in front of their computers but offered flexibility through the chatbot so that they could utilize their computers or mobile devices for engaging in EMAs, surveys, or interventions as needed.

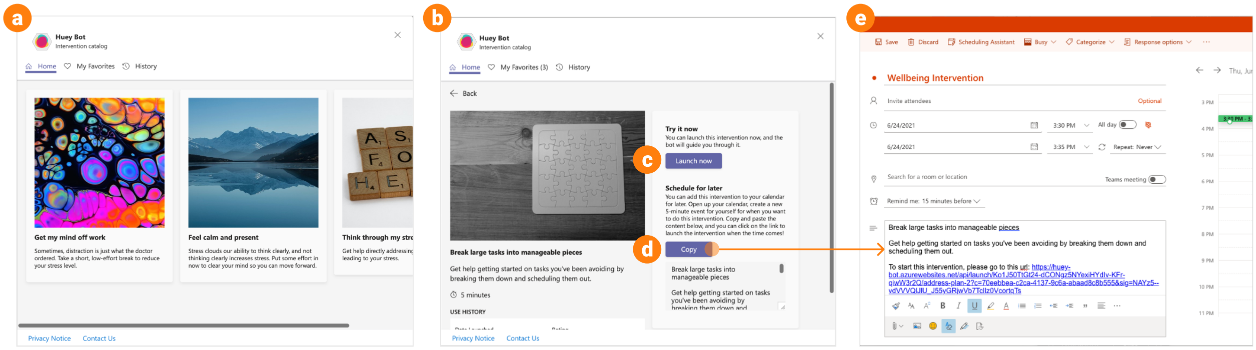
Stress Score Component

The *stress score* component is responsible for inferring the user's stress level based on passively-sensed information. Sano et al. [355] have highlighted that computer usage, calendar and email usage, intervention history, activity and heart rate variability are together useful features for predicting intervention timing. Our *stress score* component uses similar features but optimizes for detecting moments of high stress, such that interventions can be delivered at moments of need (i.e., reduce momentary high stress).

We capture contextual and behavioral information about people through custom logging software that runs on their primary work computer. From this logged information that includes computer activity (e.g., window switching, keyboard usage) and behavioral and physiological signals (e.g., facial expression, breathing rate), we compute a stress score. Specifically, the logging software has three main capabilities. The first is to capture email, calendar, and application data from the users' desktop applications usage (all participants used Microsoft Outlook as their primary email and calendar software). The second is to use their webcam to capture their position and facial actions while they are in front of their workstation. The third is to use their webcam to capture their heart rate and breathing rate using a non-contact measurement technique [246].

We designed the stress score to capture five components that previous work has identified as sources of stress. They are defined as follows:

- **Email** (w_1): The volume of emails received in a given day has been linked to higher stress in information workers [260, 261]. The *email* component (w_1) at X hours into the day was computed as the number of emails received until that time of day / 2400.
- **Calendar** (w_2): The lack of breaks and number of appointments in a work day (e.g., meetings) is a known stressor for information workers [262]. The *calendar* component (w_2) of the score was computed as the total number of appointments in a given day / 15.
- **Time** (w_3): People are also more likely to experience negative emotional states (such as stress) later in the day, in general [274]. The percentage into the day (w_3), was the *time* component of the stress score.
- **Facial** (w_4): Previous work has identified that facial expressions during information work can capture changes in affect [270, 274]. To describe facial behavior we use the Facial Action Coding System (FACS) [131], the most commonly used and descriptive taxonomy. The *facial* component (w_4) was computed as corrugator (AU04) and lip depressor (AU15) activity minus zygomatic major



Get my mind off work

Sometimes, distraction is just what the doctor ordered. Take a short, low-effort break to reduce your stress level.

Feel calm and present

Stress clouds our ability to think clearly, and not thinking clearly increases stress. Put some effort in now to clear your mind so you can move forward.

Think through my stre

Get help directly addressing leading to your stress.

Break large tasks into manageable pieces

Get help getting started on tasks you've been avoiding by breaking them down and scheduling them out.

5 minutes

USE HISTORY

Try it now

You can launch this intervention now, and the bot will guide you through it.

Schedule for later

You can add this intervention to your calendar for later. Open up your calendar, create a new 5-minute event for yourself for when you want to do this intervention. Copy and paste the content below, and you can click on the link to launch the intervention when the time comes!

Break large tasks into manageable pieces

Get help getting started on tasks you've been avoiding by

Wellbeing Intervention

Invite attendees

6/24/2021 3:30 PM All day

6/24/2021 3:35 PM Repeat: Never

Search for a room or location

Teams meeting

Remind me: 15 minutes before

Break large tasks into manageable pieces

Get help getting started on tasks you've been avoiding by breaking them down and scheduling them out.

To start this intervention, please go to this url: https://huey-bot.azurewebsites.net/api/launch/Ko1J50TtGt24-dCONgz5NYexiHYdIv-KFr-qiwW3r2Q/address-plan-2?c=70eebbea-c2ca-4137-9c6a-abaad8c8b555&sig=NAYz5--vdVVVQIJlU_J55yGRjwVb7TcIlz0VcortqTs

Figure 5.1: Users can browse the intervention catalog to choose from different intervention types (a) and navigate to an intervention to view its details (b). Users can choose to launch the intervention at that moment (c) or copy relevant information about the intervention, including a link to launch the intervention, to perform the intervention at a later time (d). Users can paste this information into a calendar event to schedule it at the desired future time (e).

(AU12). Corrugator (i.e., brow furrowing) and lip depressor (i.e., frowning) are typically linked to negative experiences (e.g., confusion, frustration, displeasure), whereas zygomatic major (i.e., smiling) is typically linked to positive experience (e.g., joy, pleasure) [209].

- **Physiological** (w_5): Changes in heart rate are also associated with increased stress during computer-based work [427, 186]. The *physiological* component of the stress score (w_5) was computed as the current heart rate (in beats/min) divided by 100 beats/min.

We normalize each component of the stress score to create a value between 0 and 1. If any of these numbers was greater than 1, it was rounded to 1. We combined and normalized these components to compute the stress score as:

$$S = \frac{w_1 + w_2 + w_3 + w_4 + w_5}{5} \tag{5.1}$$

This stress score was stored in the database for retrieval by the *JIT* component.

Our stress score includes aspects of work demand (i.e., email volume, calendar volume) and available resources (i.e., time into the day) as well as behavioral and/or physiological changes in reaction to stress (i.e., facial expression, heart rate). These components were designed based on insights from prior work and crafted to create a simple and explainable continuous estimate of how likely an individual was to be experiencing stress. Although a more elaborate stress score could be used in the future, for the purposes of our study, we found that this was a practical estimate of stress. We ran a retrospective analysis of the correlation between our stress scores and participants' self-reported ratings on their momentary stress levels via EMAs (1=Not at all stressed; 3=Moderately stressed; 5=Extremely stressed), we found a significant positive correlation between the two (N=1318, Pearson r=0.2, p<0.01). We note that the stress score was not always available at the time participants reported their momentary stress levels because the participant may have disabled the sensing software temporarily or responded to the EMAs when they are not at their desks. Therefore, we rely on both stress scores and self-reported stress levels for the *JIT* component, which we discuss next.

JIT Component

Our *JIT* component is responsible for determining if the system should nudge the user to perform a stress-reduction intervention at that time. We leverage the computed stress score and also apply heuristics that incorporate users' self-reported stress levels to maximize the potential efficacy of the stress interventions [381]. Self-reported stress levels are obtained from EMAs or at the end of intervention usage.

Based on the data from the first week of using the system during the four-week observation period, we compute each user's average for computed stress scores and self-reported stress levels. We then use these averages as individual baselines (week one of four) and as thresholds for delineating high-stress from low-stress during the subsequent weeks (weeks two to four). Our logic for JIT interventions also checks historical intervention usage and system nudges to ensure that the system does not aggressively prompt engagement. Our system will send a nudge to engage in a stress intervention if and only if:

- Computed stress score is greater than or equal to the user's baseline (or 0.5) in the past 5 minutes *or* self-reported stress level is greater than or equal to the user's baseline (or Moderately Stressed) in the past 30 minutes,
- It is during the weekday and the user's working hours,
- There are no scheduled stress interventions during the remainder of the day,
- The user has not completed an intervention in the past hour,
- There has not been a system-initiated nudge in the past two hours, and
- There have not been four or more nudges so far that day.

These nudges are sent as messages from the chatbot, which we describe next. Our system is configurable such that the *JIT* component can be active for a subset of users.

Chatbot Component

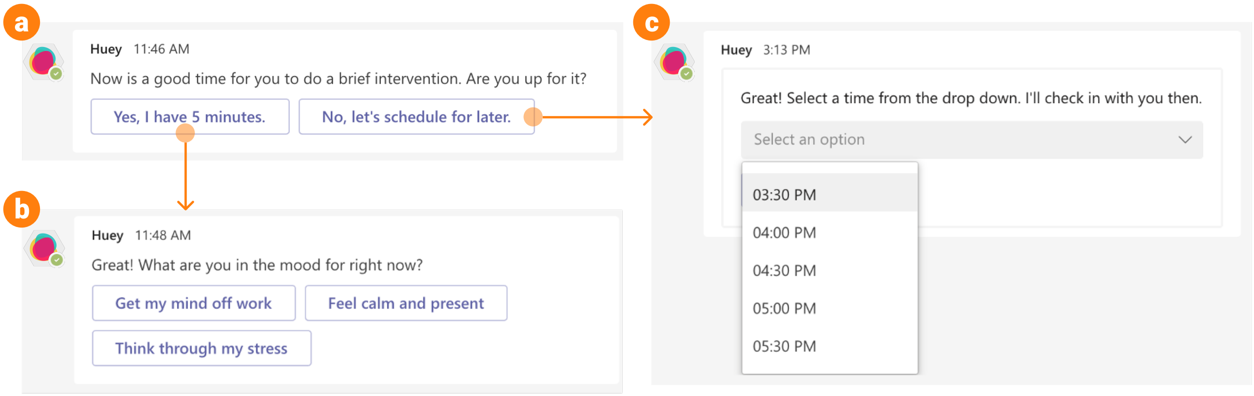
We use the Microsoft Teams chatbot[1] as a platform to maximize success in delivering and interacting with EMAs, surveys, and intervention content because all of our users regularly used Teams for work-related communication and had Teams clients readily available on desktops and mobile phones. We designed our chatbot, named Huey, to proactively initiate conversations with the users, and it would remind users to complete EMAs or surveys or to help them engage in stress-reducing interventions. Most prompts are presented as Adaptive Cards[2], either with predefined response options (e.g., scales for stress levels) or a button to open a task module dialog that hosted web-based contents (e.g., videos, surveys). We implemented Huey using Microsoft's Bot Framework[3].

---

[1] https://docs.microsoft.com/en-us/microsoftteams/platform/bots/what-are-bots

[2] https://adaptivecards.io/

[3] https://dev.botframework.com/



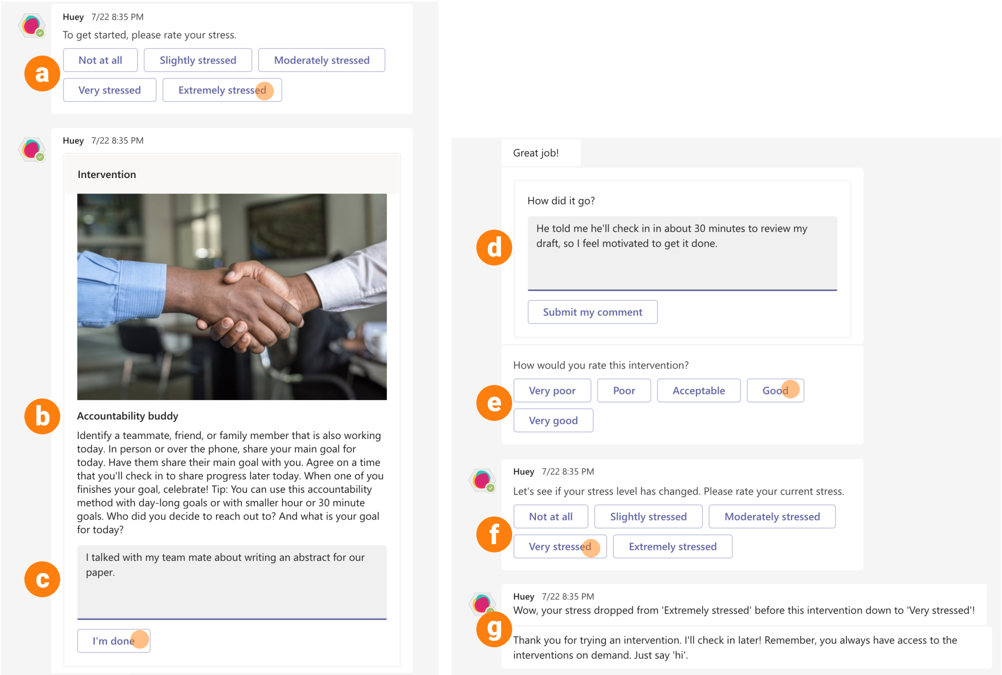
Now is a good time for you to do a brief intervention. Are you up for it?

Great! What are you in the mood for right now?

Great! Select a time from the drop down. I'll check in with you then.

Figure 5.2: When the system determines that an intervention is needed at the moment, Huey sends a nudge to users to perform an intervention (a). When the users opt to perform interventions at that moment, they can choose an intervention type they are interested in performing (b). Or they can opt to postpone the intervention and select a later time for Huey to check back in (c).

Huey provides seamlessly integrated experiences for browsing the intervention catalog or consuming the stress interventions through Teams task modules (i.e., embedded web controls) so that the users can achieve all tasks within the Teams app. Users can browse the intervention catalog to drill into different intervention types (Figure 5.1a), navigate to an intervention they like (Figure 5.1b), and launch the intervention within the same dialog flow (Figure 5.1c). If the users want to perform the intervention at a later time, they can copy intervention metadata, which includes a link to launch the intervention at any time (Figure 5.1d). This information can be easily pasted into a calendar event and scheduled at a time that works better for them (Figure 5.1e). When Huey nudges users to perform an intervention, users can choose to perform the intervention at that moment or to postpone it to a later time that day (Figure 5.2a). When the users opt to perform interventions at that moment, they can choose an intervention type they are interested in performing (Figure 5.2b). Then Huey selects a random intervention from that category that is used the least frequently. Or users can opt to postpone the intervention and select a later time for Huey to check back in (Figure 5.2c).



Huey 7/22 8:35 PM
To get started, please rate your stress.

Huey 7/22 8:35 PM
Intervention

Accountability buddy

Identify a teammate, friend, or family member that is also working today. In person or over the phone, share your main goal for today. Have them share their main goal with you. Agree on a time that you'll check in to share progress later today. When one of you finishes your goal, celebrate! Tip: You can use this accountability method with day-long goals or with smaller hour or 30 minute goals. Who did you decide to reach out to? And what is your goal for today?

I talked with my team mate about writing an abstract for our paper.

Great job!

How did it go?

He told me he'll check in in about 30 minutes to review my draft, so I feel motivated to get it done.

How would you rate this intervention?

Huey 7/22 8:35 PM
Let's see if your stress level has changed. Please rate your current stress.

Huey 7/22 8:35 PM
Wow, your stress dropped from 'Extremely stressed' before this intervention down to 'Very stressed'!

Thank you for trying an intervention. I'll check in later! Remember, you always have access to the interventions on demand. Just say 'hi'.

Figure 5.3: An example dialog flow of how Huey facilitates intervention consumption. Huey first asks users to rate their current stress level (a) before presenting the intervention content to the users (b). The intervention may ask the users to respond to a prompt (c). After completing the intervention, Huey asks users to reflect on the intervention (d) and provide a rating (e). Finally, Huey asks users to rate their stress level (f) and comments on the changes in stress levels (g).

Just before the users perform their intervention, Huey asks them to rate their momentary stress using a 5-level stress rating (1=Not at all stressed; 3=Moderately stressed; 5=Extremely stressed; Figure 5.3a). After performing the intervention, Huey asks users to reflect on how the intervention went and to rate the intervention (1=Very poor; 3=Acceptable; 5=Very good; Figure 5.3d,e). Finally, Huey concludes the intervention by asking users to rate their momentary stress and comments on the changes in stress levels from before performing the intervention (Figure 5.3f,g). Figure 5.3 shows an example of this intervention consumption flow.

Huey supports three different intervention modalities: (1) A video-based intervention provides a brief description of the content, followed by a task module dialog that played the video (Figure 5.4a); (2) A single-turn text prompt intervention provides a brief instruction for the users to engage in an activity, followed by a prompt to answer a reflective question (Figure 5.4b); (3) A conversation-based intervention provides a dialog that guides users through a series of prompts (Figure 5.4c).

Intervention Design

We designed interventions based on components of Cognitive Behavioral Therapy (CBT) and Dialectical Behavioral Therapy (DBT), two empirically supported front-line psychotherapy modalities that are used to flexibly treat a wide range of mental health and well-being concerns [73, 251]. We created all interventions to take under five minutes, and they were comprised of either a short video, a single-turn text prompt, or a brief therapeutic conversation with Huey. We then categorized the interventions depending on the function served for users as well as the approximate effort required.

- **Get my mind off work** (Low effort): ‘Get my mind off work’ interventions are translations of the DBT *pleasant activities schedule*, a list of positive activities that help individuals regulate their emotions by becoming engaged with something that elicits positive feelings [245]. Examples include watching a short video of penguins playing in Antarctica or listening to a favorite song. These interventions are simple and likely similar to activities many individuals do naturally throughout the day when attempting to take a break at work. However, they steer clear of activities that may feel pleasant in the moment but research shows may lead to more distress in the long-term, such as scrolling on social media or eating a high quantity of sugary food. These interventions are intended to capture users’ attention with low levels of user effort and investment and are deployed using the single-turn text prompt and video formats.
- **Feel calm and present** (Medium effort): ‘Feel calm and present’ interventions draw inspiration from the mindfulness practices taught in CBT and DBT, which

help individuals re-focus on the present moment in order to gain perspective and increase control over their thoughts, feelings, and behaviors. Examples include using the five senses to notice and name components of the immediate environment or writing a self-affirming statement ten times with one's non-dominant hand – an activity that takes a substantial amount of focus and tends to bring individuals in states of stress into the present moment. These interventions typically require moderate levels of user effort but have significant empirical support as stand-alone interventions capable of decreasing stress [245]. They are deployed using single-turn text prompt and video formats.

- **Think through my stress** (High effort): 'Think through my stress' interventions help users directly address and problem-solve stress-inducing components of their lives using strategies from CBT and DBT, such as cognitive re-framing, pros and cons lists, and reaching out to a friend or co-worker for help with emotional processing or getting productive [441]. These interventions require the highest amount of user effort as they necessitate direct engagement with stressful content. They are delivered via single-turn text prompts and therapeutic conversations with Huey.

Study Design

We conducted a four-week, between-subjects user study, where our participants engaged with our system through the Huey chatbot, which delivered stress-reducing micro-intervention content and facilitated different study requirements and protocols.

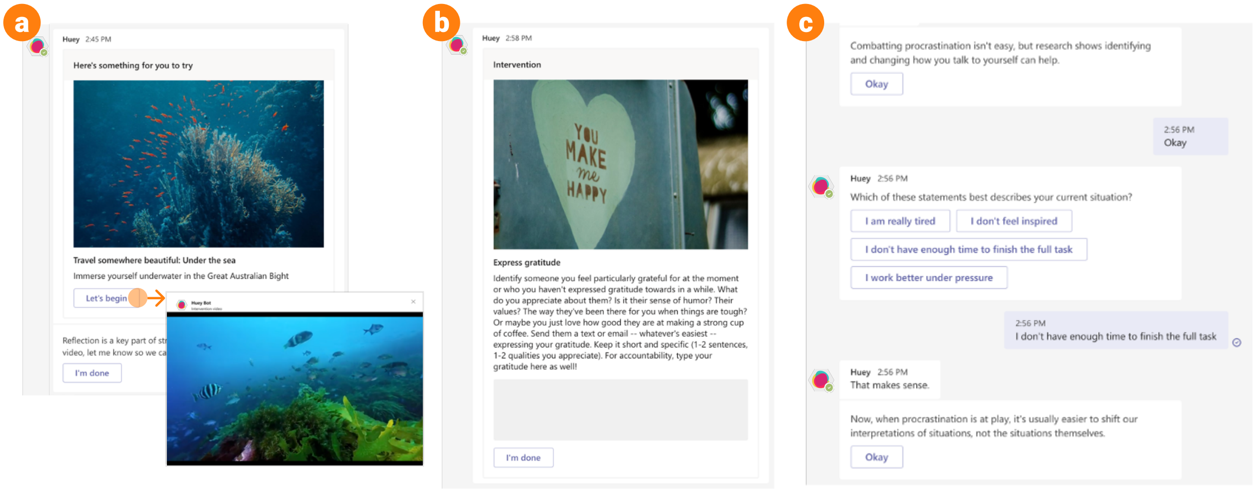
Participants

We recruited information workers from a large technology organization by sending email advertisements about the study to a randomly sampled set of employees from the organization's employee database. Interested participants completed a brief screener survey about their demographics (e.g., age, gender, role) and work set up (e.g., primary device specification and OS, web camera availability). Eligible participants, whose primary device specifications met our sensing software requirements, were asked to install and run the study software for 30 minutes to confirm system compatibility. We then enrolled participants on a first-come, first-served basis. In total, we enrolled 87 participants in the study. Participants were randomly assigned to one of the two conditions while maintaining equal gender distribution between the conditions, as prior work shows women report higher overall workplace stress than men [184]. One participant dropped out and another participant switched conditions during week one of the four-week observation period, both due to unforeseen technical issues.

Of the final set of 86 participants who successfully completed the study, 65.1% identified as male and 32.6% as female. 38.4% were in the age range of 36-45 years old, 23.3% in 26-35, and 23.3% in 46-55 years old ranges. 54.7% worked in Engineering/Development roles, 22.1% in Sales and Marketing roles, 8.1% in Operations and Services roles, 5.8% in Business Development and Strategy roles, and 4.7% in Administrative Assistant or Human Resources roles. 86.2% of the participants worked remotely from home.

Huey also supports special commands to allow accessing study-related instructions (via "help") or on-demand interventions (via "hi") at any time. Messaging "hi" to Huey would initiate dialogs for browsing the intervention catalog and performing interventions on demand. The timing of reminders for EMAs and surveys as well as which dialog flows are available are configurable per user.

We were unable to capture webcam-based signals from 16 of 86 participants (eight in each condition) due to unforeseen performance issues with the sensing software that interfered with their daily work. However, our *stress score* component is robust to missing data such that this was not an issue for the study.



Huey 2:45 PM

Here's something for you to try

Travel somewhere beautiful: Under the sea

Immerse yourself underwater in the Great Australian Bight

Reflection is a key part of str

video, let me know so we ca

Huey 2:58 PM

Intervention

Express gratitude

Identify someone you feel particularly grateful for at the moment or who you haven't expressed gratitude towards in a while. What do you appreciate about them? Is it their sense of humor? Their values? The way they've been there for you when things are tough? Or maybe you just love how good they are at making a strong cup of coffee. Send them a text or email -- whatever's easiest -- expressing your gratitude. Keep it short and specific (1-2 sentences, 1-2 qualities you appreciate). For accountability, type your gratitude here as well!

Combatting procrastination isn't easy, but research shows identifying and changing how you talk to yourself can help.

2:56 PM
Okay

Huey 2:56 PM
Which of these statements best describes your current situation?

2:56 PM
I don't have enough time to finish the full task

Huey 2:56 PM
That makes sense.

Now, when procrastination is at play, it's usually easier to shift our interpretations of situations, not the situations themselves.

Figure 5.4: Huey supports three intervention modalities: (a) A video-based intervention provides a brief description of the content, followed by a task module dialog that played the video; (b) A single-turn text prompt intervention provides a brief instruction for the users to engage in an activity, with an open-ended prompt to answer a reflective question; (c) A conversation-based intervention provided a dialog that guides engage through a series of prompts.

Procedure

The study procedure included one week of onboarding, four weeks of observation of intervention usage and engagement, and one week of off-boarding (Figure 5.5).

During the preparation week, participants were asked to install the required sensing software and the chatbot and to complete an intake survey. The intake survey asked for the participant's local time zone and their typical start and end times of the work day, which were used to configure the system's interaction with the participants. The intake survey included the Depression, Anxiety, and Stress Scale 21 (DASS-21), a 21-item self-report questionnaire designed to assess clinical levels of depression, anxiety, and stress [248], and the Emotional Regulation Questionnaire (ERQ), a 10-item scale designed to measure respondents' tendency to regulate their emotions through cognitive reappraisal and expressive suppression [161]. We asked participants to report their current stage of behavior change to reduce stress at work from four stages of behavior change (Stage 1: Pre-contemplation, Stage 2: Contemplation, Stage 3: Taking action, Stage 4: Maintenance) adapted from the Transtheoretical Model [320] The intake survey also included questions about personality style, stressful life events, emotional resilience, and self-care style.

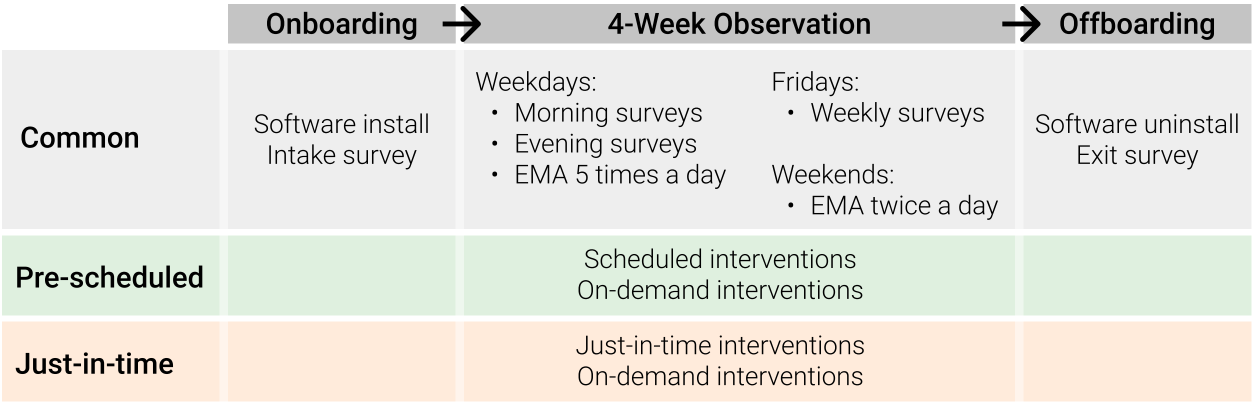
During the four-week observation period, participants were asked to interact with Huey to engage in stress-reducing interventions, where we configured the intervention system to enable features specific to their assigned conditions:

- **Pre-scheduled Engagement (PS)**: In this condition, participants were asked to plan their interventions in advance. On Fridays prior to each study week, Huey asked participants to browse the catalog of stress-reducing interventions, choose specific interventions they would like to try, and schedule at least one intervention for the upcoming week into their work calendar (Figure 5.1). From the intervention catalog, participants could copy intervention-specific information into their calendars with a link to launch the intervention. These participants leveraged the calendar's built-in reminder functionality for the interventions. On Mondays of each study week, Huey reminded participants to review their scheduled interventions and to adjust them appropriately. When the scheduled time arrived, participants clicked on the link in the calendar event to engage with Huey to carry out the intervention. Participants in this condition could also access the intervention catalog on demand, where they could launch a selected intervention at that moment or copy intervention-specific information into their calendars.
- **Just-in-time Engagement (JIT)**: In this condition, participants were asked to engage with an intervention based on our system's JIT component (Section 5.1.1). When our JIT component determined that a stress-reducing intervention is needed and appropriate for the participant, Huey sent a message to the participant, presenting an option to perform the intervention at that moment or

> to postpone it to a later time that day (Figure 5.2). When the participants opted to perform interventions at that moment, Huey asked them to choose between one of the three intervention categories we described in Section 5.1.1. Once a category was selected, Huey chose a random intervention from that category that was used the least frequently. Participants then engaged with Huey to carry out the intervention. As in the PS condition, participants in the JIT condition could also access interventions on demand, where they could perform the intervention at that moment or schedule it for a later time that day.

Based on each participant's reported work hours, Huey asked participants to complete five EMAs per day during the weekday, roughly equally spaced apart to be on the hour or half-hour (e.g., at 9 AM, 11 AM, 12:30 PM, 2:30 PM, and 4 PM if work hours span 9 AM to 5 PM), and to complete two optional EMAs during the weekends (e.g., at 11 AM and 3 PM). Each EMA consisted of two parts. The first part was required for the study and asked participants to rate their stress level during the past 30 minutes using the same 5-level stress rating. The second part included questions about work demands, available resources, arousal, valence, food consumption, and social interactions. EMA questions can be found in the Supplementary Information. Participants were also asked to complete morning surveys 15 minutes before the start of each work day, evening surveys 15 minutes before the end of each work day, and weekly surveys during the afternoon on Fridays. The morning survey included questions from the Census Sleep Diary [76], and the evening survey captured food and beverage intake throughout the day. On Fridays of the first three weeks, participants were asked to complete weekly surveys which included questions from the DASS-21 and the stages of behavior change.

After the four-week observation period, participants were asked to complete an exit survey that included scales for DASS-21, stressful life events, emotional resilience, and the stages of behavior change. The exit survey presented 8 questions probing



	Onboarding	4-Week Observation	Offboarding
Common	Software install Intake survey	Weekdays: • Morning surveys • Evening surveys • EMA 5 times a day Fridays: • Weekly surveys Weekends: • EMA twice a day	Software uninstall Exit survey
Pre-scheduled		Scheduled interventions On-demand interventions	
Just-in-time		Just-in-time interventions On-demand interventions	

Figure 5.5: Our study procedure included onboarding with software install and intake survey, four weeks of observation of intervention usage, surveys, and EMAs, and off-boarding with software uninstall and exit survey.

the usability of the assigned conditions including ease of use, satisfaction, and frustration. The exit survey also included condition-specific and open-ended questions probing their preferences for engaging with the interventions, appropriate timing of interventions, and how participants compared accessing interventions on-demand, scheduling interventions in advance, or being nudged to do an intervention by a system. It also solicited comments about the intervention content, what factors motivated them to perform the interventions, and any perceived helpfulness or impact of the interventions on stress reduction. These questions were identical for both conditions. Exit survey questions can be found in Appendix C.

Each participant was compensated with a $400 Amazon gift card for their participation and data. Our study was approved by the Microsoft Research Institutional Review Board (IRB).

Data Processing & Analysis

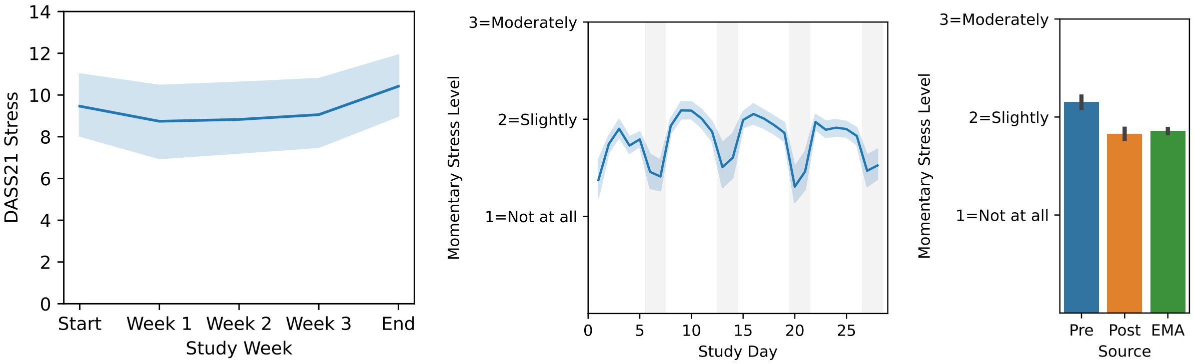
We combined data from the system usage logs and survey responses to understand engagement patterns, intervention usage, and outcomes. We leveraged the system usage logs for our analysis of intervention usage. Each intervention use instance was associated with one of three intervention categories (Section 5.1.1), whether or not it was used on-demand, timestamps of when it was started and completed, stress levels immediately before and after the intervention use, user ratings, and any free-form comments the participants wanted to provide. We extracted 1651 unique intervention attempts during our study. 28.9% (477/1651) of those were started but never completed. 96.5% (1133/1174) of the completed interventions were followed by user ratings, and 92.4% (1085/1174) of the completed interventions had both pre-intervention and post-intervention stress levels. We collected 6685 stress levels from EMAs, 1174 from pre-intervention use, and 1085 from post-intervention use, for a total of 8944 stress levels. We had both intake and exit DASS-21 measures for each of the 86 participants and 217 DASS-21 measures from weekly surveys, for a total of 389 DASS-21 measures. We computed *momentary stress reduction* by subtracting the pre-intervention stress levels from the post-intervention stress levels and *study-long stress reduction* by subtracting the DASS-21 stress sub-scale responses from the intake surveys from that of the exit surveys, where positive values indicate higher stress reduction. We aggregated the intervention usage and stress level data for each participant for analysis. We mapped the participants' reported stages of behavior change to numerical values based on their reported stage (Stage 1: Pre-contemplation, Stage 2: Contemplation, Stage 3: Taking action, Stage 4: Maintenance) and examined the *change in the stages of behavior change* between the study start and end.

For comparing the means of the two conditions (JIT vs. Pre-scheduled), we used the Welch Two Sample t-test. Wherever applicable, we used Benjamini-Hochberg procedure [41] on the t-test results to correct for multiple comparisons. For comparing differences among participants, we used paired t-tests. We used one-way analysis of variance (ANOVA) tests to examine differences in outcome variables (e.g., stress reduction) with multi-level factors (e.g., intervention categories). When we found significant results, we then investigated pairwise differences, employing Tukey's HSD procedure to correct for the increased risk of Type I error due to unplanned comparisons. We used linear mixed-effects models to investigate the relationship between per-participant characteristics and outcome variables, again investigating any pairwise differences using Tukey's HSD procedure to adjust for repeated testing. When we included gender as a variable, we included the subset of participants (N=84) who identified themselves as male or female, due to the small sample size of other gender identity categories (N=2). For correlation analyses, we used Pearson's correlation. We used Python and R for processing the data and for statistical analyses.

Two researchers qualitatively coded the open-ended survey responses using inductive thematic analysis [57]. We identified several topics of interest (e.g., the timing and frequency of engagement with the bot, motivating factors, preferences for engagement and interventions, desired functionalities), categorized participant responses into themes within each topic, and quantified their occurrence.

5.1.2 Findings

We first describe the temporal trajectory of participants' self-reported stress throughout the study to contextualize the overall impact of the study (Section 5.1.2). Then we organize our findings according to our research questions. We first address RQ1, presenting the impact of the two engagement timing conditions on the overall intervention usage, momentary and study-long stress reduction, and user ratings (Section 5.1.2). We also include the impact of on-demand intervention usage. We then address RQ2 and present the impact of the three intervention types on usage, stress reduction, and ratings (Section 5.1.2). Finally, we address RQ3 and



(a) Weekly DASS21 stress level

(b) Daily momentary stress level

(c) Momentary stress level across sources

Figure 5.6: Self-reported stress levels throughout the study. (a) shows the average DASS-21 stress subscale from intake, weekly, and exit surveys. DASS-21 stress scores of 0-14 indicate a 'Normal' stress range. (b) shows average momentary stress levels per day from EMAs and pre-/post-intervention uses. Vertical grey bands denote weekends to highlight daily patterns on momentary stress levels. (c) shows average momentary stress levels across all pre-intervention uses (Pre), post-intervention uses (Post), and EMAs. All error bands and bars indicate 95% confidence intervals.

summarize participants' feedback on overall system usability, engagement timing, and interventions (Section 5.1.2).

Stress over Time

Study-long stress Overall, the present sample did not experience extreme levels of stress or statistically significant changes in stress levels from study start to study finish. Average participant stress at study start was within a non-clinical range (i.e., within normal limits; $\bar{x}$=9.47, σ=6.96 at study start) and remained within this range throughout the study with little variation (Figure 5.6a), as measured by the stress sub-scale of DASS-21 [248]. Stress levels at the study start and stress levels at the study end ($\bar{x}$=10.42, σ=7.23) indicated no statistically significant study-long change in stress levels.

Short-term stress Momentary stress levels, as assessed via EMAs and pre-/post-intervention stress levels, were relatively stable throughout the course of the study as well. Participants reported an average of 104 momentary stress levels over the course of the four-week observation ($\bar{x}$=4.19, σ=2.39 per day per participant). The average momentary stress level was 1.87 (σ=0.91), between '1=Not at all stressed' and '2=Slightly stressed,' with minimal variation over the course of the study. We also saw lower stress levels during the weekends (Figure 5.6b). Average EMA stress level was generally below pre-intervention and above post-intervention stress levels (Figure 5.6c).

Pre-/post-intervention stress Pre-intervention ($\bar{x}$=2.16, σ=0.95) and post-intervention ($\bar{x}$=1.82, σ=0.87) stress levels indicate that interventions resulted in a statistically significant momentary decrease in stress levels ($t(1084)=18.113$, $p \ll 0.001$).

RQ1: Engagement Timing Impact

Quantity of interventions used by timing Participants completed an average of 13.65 interventions over the four-week study ($\bar{x}$=10.39, min=2, max=55). We found a statistically significant difference in number of interventions completed between the two conditions; JIT participants completed significantly more interventions than PS participants (19.74 vs. 7.56 per participant, $t(63.633)=-6.696$, $p \ll 0.001$), but this is likely due to the study design, wherein JIT participants were prompted throughout the day.

Stress reduction by timing Despite differences in the completed usage of interventions, we found no statistically significant difference in both momentary and study-long stress reduction between the two conditions. We also found no correlation between the total number of completed interventions and study-long stress reduction (Pearson r=-0.06).

User ratings by timing Participants had a generally positive reaction to the interventions, giving them an average rating of 3.65, between '3=Acceptable' and '4=Good' (σ=0.98). We found that JIT participants rated interventions significantly lower than PS participants by about 0.256 points ($\chi^2(1)$=5.962, $p<0.05$).

Behavior change stage by timing At the beginning of the study, 48.8% of participants were in the 'Stage 3: Taking action' stage of behavior change, with 32.6% in 'Stage 2: Contemplation', 15.1% in 'Stage 4: Maintenance', and 3.5% in 'Stage 1: Pre-contemplation' stages. Controlling for the behavior change stage at the study start, we found a statistically significant difference in advancement through the behavior change stages between conditions: PS participants reported significantly more advancement through the behavior change stages compared to JIT participants ($F(1)$=6.834, $p<0.05$) and no statistically significant interaction effect between the intake stage and condition.

On-demand usage by timing Although users could access on-demand interventions in both conditions, PS participants completed statistically significantly more on-demand interventions compared to JIT participants (2.63 vs. 0.02; $t(42.125)$=7.552, $p \ll 0.001$). On average, PS participants completed interventions on-demand 38.2% of the time (σ=0.33). 46.5% of PS participants completed on-demand 50% of the time or more. For PS participants, interventions used on demand reduced statistically significantly more stress than those used at pre-scheduled times ($\chi^2(1)$=10.587, $p<0.01$) by about 0.23 points. According to their pre-intervention stress levels, we found that PS participants used on-demand interventions when they were slightly more stressed than at pre-scheduled times (2.23 vs 2.06), but the effects were not statistically significant ($t(229.81)=-1.862$, $p=0.064$). We found no statistically significant difference in subjective ratings between interventions used on-demand versus those used at pre-scheduled times.

RQ2: Intervention Type Impact

In both PS and JIT conditions, participants were able to choose from the three intervention types whenever completing an intervention.

Quantity of interventions used by type On average, participants selected 'Get my mind off work' interventions 36.8% ($\sigma=0.254$) of the time and completed 71.6% of those selected. They selected 'Feel calm and present' 45.5% ($\sigma=0.251$) of the time and completed 72.7% of those selected, and selected 'Think through my stress' 17.7% ($\sigma=0.201$) of the time and completed 100% of those selected. JIT participants completed statistically significantly more 'Feel calm and present' interventions ($t(82.631)=-2.978$, $p<0.01$) and statistically significantly fewer 'Get my mind off work' interventions ($t(80.364)=-2.073$, $p<0.05$) compared to PS participants. There were no statistically significant usage differences between JIT and PS for the 'Think through my stress' intervention type. We modeled the impact of baseline DASS-21 stress, emotion regulation strategies, behavior change stage, age, and gender on the completion rate per intervention type and found the baseline DASS-21 stress to have a statistically significant effect on the completion rate of 'Feel calm and present' interventions ($F(1)=5.630$, $p<0.05$). We found no other statistically significant effects.

Stress reduction by type When we examined the impact of the completion rate per intervention type on study-long stress reduction, we found that a higher rate of completed uses of 'Get my mind off work' among all completed uses had a statistically significant improvement on study-long stress reduction ($F(1,83)=6.055$, $p<0.05$). Of

1085 completed intervention uses with pre- and post-intervention stress levels, ‘Think through my stress’ reduced momentary stress by 0.41 points on average (σ=0.61), ‘Get my mind off work’ reduced momentary stress by 0.32 points on average (σ=0.55), and ‘Feel calm and present’ reduced momentary stress by 0.26 points on average (σ=0.53; Figure 5.7a). Intervention type had a statistically significant effect on momentary stress reduction ($\chi^2(1)$=9.77, p<0.01). Pairwise comparisons of intervention type revealed that ‘Get my mind off work’ interventions reduced momentary stress more than ‘Feel calm and present’ interventions, and ‘Think through my stress’ interventions reduced momentary stress more than ‘Feel calm and present’–both to a statistically significant extent. There was no statistically significant difference in momentary stress reduction between ‘Think through my stress’ and ‘Get my mind off work.’ We found similar results when controlling for condition (JIT/PS), baseline stress, emotion regulation style, behavior change stage, gender, and age.

User ratings by type On average, participants’ subjective ratings of the interventions were 3.67, between ‘3=Acceptable’ and ‘4=Good,’ for ‘Get my mind off work’ (σ=1.02), 3.67 for ‘Feel calm and present’ (σ=0.96), and 3.52 for ‘Think through my stress’ (σ=0.96; Figure 5.7b). The intervention category had a statistically significant effect on user rating ($\chi^2(1)$=7.44, p<0.05), such that ‘Get your mind off stress’ interventions were rated significantly higher than ‘Think through my stress’ interventions.

RQ3: User Feedback on Intervention Timing and Types

Pre-scheduled participant feedback PS participants (N=43) used a variety of factors for determining when they would choose to place interventions on their calendars. Some participants chose the beginning of the day or the end of the day (N=24), and several chose to space them throughout the week (N=7). Many looked for

(a) Momentary stress level across intervention types

(b) User rating across intervention types

Figure 5.7: Average momentary stress reduction (a) and user rating (b) across three intervention types. Positive stress reduction indicates that stress is reduced after intervention use on a 5-point scale (1=Not at all stressed; 3=Moderately stressed; 5=Extremely stressed). Higher user rating on a 5-point scale (1=Very poor; 3=Acceptable; 5=Very good) indicates that the user liked the intervention after intervention use. The error bars indicate 95% confidence intervals across all intervention uses.

free spots on their calendar (N=13), after several back-to-back meetings when they knew they would be stressed, or afternoons when they knew they would be tired.

PS participants liked that having interventions on the calendar held them accountable (N=13): "I didn't forget because it was on the calendar"; "it calmed me seeing it was there." There was a subgroup (N=13) that especially liked to plan interventions out or make them recur, while others specifically mentioned their ease of use (N=5), that they could use them on demand if they needed to (N=3), and that they liked having breaks in the calendar to learn something new (N=6). Although some users complained that the chatbot did not schedule interventions automatically for them based on their availability, others noted that free times were not necessarily indicative of stressful moments, meaning a chatbot might not pick the best time based on a simple free time algorithm.

While 11 out of 43 PS participants told us that they liked pre-scheduled and on-demand interventions equally, 30 participants stated a strong preference for on-demand interventions because they were easy to access and perform when they were stressed (in the moment) and because it was too hard to predict when they would be stressed in the future. Though on-demand interventions were easy to access, it was noted that it was hard to remember to do them. 33 participants said that they wanted automatic “nudging” by the bot based on their stress levels.

JIT participant feedback JIT participants (N=43) thought that the JIT interventions were a good reminder to take time out of their day, especially when stressed (N=30). They thought that the interventions were convenient and helpful (N=17). In terms of improvements to the design of the system, JIT participants raised timing and frequency issues: nudges were too frequent and disruptive of focus (they thought that the system should make them easier to ignore). Although they wanted automatic detection and interventions based on stress levels and an overly crowded calendar, they suggested intelligent timing based on their availability and task context. Many JIT participants liked the agency to perform the interventions when they wanted or needed to by using the on-demand feature, which ended up being less disruptive to workflow and kept the users in control (N=18).

System feedback On a 5-point agreement scale (1=Strongly disagree; 5=Strongly agree), participants in both conditions agreed that the intervention system they used during the study made it easier to engage in interventions compared to before the study ($\bar{x}$=4.07, σ=0.98), that the system was easy to use ($\bar{x}$=4.03, σ=1.13), that they found themselves engaging in more interventions compared to before the study ($\bar{x}$=3.88, σ=1.12), and that the system met their requirements for engaging in interventions ($\bar{x}$=3.76, σ=1.09). Participants in both conditions also agreed that, if given the opportunity, they would continue to use the system ($\bar{x}$=3.38, σ=1.20), but this had the lowest agreement scale among positive usability statements. Participants disagreed that using the system was a frustrating experience ($\bar{x}$=2.37, σ=1.18). JIT participants found their condition-specific system to be easier to use than PS participants with statistical significance (4.28 vs. 3.79, p<0.05). We found the differences between

Figure 5.8: Average agreement on eight statements about the system usability across *Pre-scheduled* (PS) condition in blue and *Just-in-time* (JIT) condition in orange. The agreement was measured using a 5-point scale (1=Strongly disagree; 5=Strongly agree). The error bars indicate 95% confidence intervals across participants. * indicates that there is a statistically significant difference between the two conditions.

conditions to be not statistically significant in all other usability ratings. Condition-specific agreement ratings are illustrated in Figure 5.8.

Intervention type feedback Generally, participants had the most positive reaction to ’Feel calm and present’ and ’Get my mind off work’ interventions. Across all participants, 30 participants felt that ‘Feel calm and present’ interventions were most helpful and 26 felt that ‘Get my mind off work’ interventions were most helpful in immediate stress reduction; only 4 participants felt that ‘Think through my stress’ interventions were helpful in immediate stress reduction. 16 and 12 participants felt that ‘Feel calm and present’ and ‘Get my mind off work’ interventions had the biggest impact on long-term stress reduction, respectively, and only 9 participants felt that ‘Think through my stress’ interventions had the biggest impact on long-term stress reduction.

Some participant reactions to specific intervention content were polarized, according to participant open-response feedback: one participant perceived ‘Think through my stress' to be "most helpful in channeling [their] energy in a new direction" while another thought that interventions that required introspection, such as those in the 'Think through my stress' category, to be least helpful because "then [they were] stressed about what [they were] writing."; one participant perceived watching videos of nature or an interesting place were "good for getting mind off current activities of the day that were the contributors to stress" while another thought "the ones which required that [they] sit at [their] computer to watch a video" to be least helpful.

Participants indicated an interest in accessing a wide variety of interventions, expressing an overall preference for interventions that varied in their physical environment (e.g., being physically away from the desk vs. doing interventions at the desk), in the level of focus on stress (e.g., think about stress vs. take the mind off stress), in the social interactions (e.g., involved other people vs. alone), in familiarity (e.g., surprising and new vs. known and expected), and in effectiveness (e.g., interventions that I benefited from). Participants overwhelmingly wanted intervention content that was simple, easy to do, and required low effort or burden.

5.1.3 Discussion

In the four-week, between-subjects study presented herein, we examined the impact of digital micro-intervention *delivery timing* and *content type* on usage patterns and stress reduction throughout the workday for N=86 information workers. Through testing two delivery timing conditions (PS and JIT) and three categories of intervention content type, we showed that digital micro-interventions were effective at reducing momentary stress (stress change from pre- to post-intervention), regardless of intervention engagement timing and content type. Although delivery timing did not have a statistically significant impact on momentary or study-long stress reduction, we did observe noteworthy differences between user perceptions of delivery timing conditions: JIT intervention engagement was perceived as easy to do and motivating, while PS intervention engagement was perceived as nicely customizable to user work schedules. We also found that PS intervention engagement was associated with statistically greater advancement through the stages of behavior change (i.e., advancement towards long-term integration of stress-reduction behaviors into everyday life). Lastly, we found that while low-effort, positive distraction interventions were perceived to be more helpful, high-effort, problem-solving interventions were indeed more effective. Understood alongside our qualitative findings, which suggested user preferences for engagement timing and content type are versatile and wide-ranging both between people and within individuals over time, we propose stress reduction intervention systems should support both PS and JIT intervention use and offer a wide variety of content types.

Design Considerations and Research Needs for Workplace Intervention Systems

Extrapolating from our quantitative and qualitative findings, we offer key design suggestions and opportunities for workplace stress-reduction intervention systems.

Integrate digital micro-interventions into the workplace.

The present study revealed that using digital micro-interventions throughout the workday reduces momentary stress and that this effect is present across variations in delivery timing (JIT, PS, and on-demand) and content type (low- to high-effort). Digital micro-intervention systems will continue to be optimized for the greatest stress-reduction impact, but even without such refinement, simply having easy access to such interventions may empower employees to reduce their stress levels within a matter of minutes. Further, digital micro-interventions were shown to reduce workplace stress when used as stand-alone individual-level interventions, i.e., as secondary strategies aimed at targeting the individual to reduce stress, without employing primary strategies at the organizational level aimed at eliminating the stressors themselves [98, 372, 332, 333]. Based on these findings, we propose that workplaces provide digital micro-intervention access to their employees as a first-line secondary strategy for reducing employee stress when employers cannot engage with primary strategies or as short-term harm reduction when primary strategies, such as changing fundamental components of an organizational structure, take months or even years to enact. Future research should compare the effectiveness of digital micro-interventions to other secondary strategies, such as longer breaks during the workday. Future research should also investigate whether certain primary strategies, such as overall workload reduction, could multiply the stress-reducing effects of digital micro-interventions and whether digital micro-intervention access could be used as a tertiary strategy for helping employees recover from stress-related mental and physical health issues.

Provide a personalized balance between automation and agency.

Our findings suggest that digital micro-intervention systems should offer users multiple levels of control over the timing and content of interventions, from low-control/high-automation options to high-control/low-automation. The majority of participants across both conditions preferred having intervention timing determined by the automated stress detection system (JIT) for ease of use – an opinion based either on lived experience from being assigned to the JIT condition or on reading a description of the JIT condition after having completed the study in the PS condition. Yet participants also requested concurrent access to interventions on-demand, the ability to pre-schedule interventions at their discretion, and the ability to “snooze” the entire system. Our findings also revealed that the JIT system tested was not sufficiently intelligent for some users due to issues like receiving intervention nudges while busy. Further, participants in the PS condition reported more advancement through the stages of behavior change over the course of the study, compared to those in the JIT condition, with the majority of participants who advanced shifting from Stage 2: Contemplation to Stage 3: Taking action. In other words, our results suggest that, despite JIT being the preferred condition, participating in the PS condition may have shifted users' self-perceptions towards being individuals capable of taking action, while participating in the JIT condition did not change users' self-perceptions. Overall, despite user preferences for JIT interventions and promises of intelligent adaptability and personalization of JITAI systems [292, 355, 386], there were benefits to user-initiated on-demand and pre-scheduled options, especially while JIT system metrics are undergoing refinement. Future research should systematically test various ratios of system automation versus user control and seek to establish whether (a) user-initiated intervention engagement promotes greater advancement through behavior change stages than future iterations of JIT systems with more sophisticated timing algorithms, and (b) which type of intervention engagement – user- versus system-initiated – is the best match for each stage of behavior change [320].

Our study also revealed a parallel user interest in system-selected content. Specifically, participants wanted to be provided with the "right" intervention for the given moment, i.e., an intervention they would *like* and that would address their momentary needs. Participants also indicated an interest in accessing a wide variety of interventions, suggesting that novelty in and of itself may be an important component of user engagement and, secondarily, intervention impact. Systems delivering digital micro-interventions should have the ability to intelligently select interventions for users depending on their momentary needs, including the need for novelty. Future research should test the frequency with which new content should be introduced and implement a content-renewal system at an optimal frequency. These features will likely lead to more sustained user engagement and stress reduction over time.

Promote self-experimentation on intervention content that compares effectiveness and effort.

Although incorporating individual preferences has been shown to improve the engagement and outcomes of stress interventions [218, 305, 365, 121], our findings suggest that users may not always be aware of which content helps reduce their stress the most. In our study, the highest-effort interventions we tested (‘Think through my stress’, i.e. problem solving) had the largest stress-reducing effect over the course of the full study, across conditions and baseline characteristics. However, these interventions were selected significantly less frequently than ‘Feel calm and present’ interventions and were rated significantly lower than ’Get my mind off work’ interventions. Future systems should offer users feedback and the opportunity to reflect on their past experiences of stress reduction and content ratings, as this may prompt different and potentially more effective choices when selecting intervention content. For example, users could benefit from a dashboard that summarizes recent trends in self-reported and passively-sensed stress levels, as well as intervention use, impact, and the rating history. By exploring past behavior, users could learn about themselves, prompting them to make more informed decisions when choosing interventions in the future.

Solicit user feedback to adapt intervention timing and content.

Just as the system providing feedback to users may facilitate their change and growth, users providing feedback to the system can help the system improve. Participant feedback in our study indicated that users are eager to provide suggestions and believe it will improve their user experience. Given the wide range of participant preferences for timing and content and the overwhelming need for personalization, this hypothesis is likely correct. Offering users intervention timing and frequency that is “just right” and providing users with personalized content for each user/context pairing will require quite sophisticated system intelligence. Opportunities for the user to train the algorithms to perform optimally will also be necessary. For example, with the help of the aforementioned user dashboard that summarizes trends in stress levels and intervention behavior, users could review their recent activity to identify patterns the system may not otherwise detect. Future research should design and test dynamic assessment and integration of user preferences into sensing and intervention delivery systems.

Limitations

The present study and associated findings have some important limitations.

Our sample presented with low levels of stress at the beginning of the study. Users of all stress levels can benefit from stress reduction interventions, but different stress levels may present different challenges for intervention design. For example, there may be more room for improvement for high-stress individuals, but high-stress individuals may be more resistant to interventions. As such, interventions should be designed for and tested on a higher-stress sample. Additionally, the relative homogeneity of the sample (all information workers, majority engineers, and majority male-identifying) limits the generalizability of our findings.

Both JIT and PS participants had access to interventions on-demand, complicating comparisons between the two conditions. Further, PS participants selected their intervention content days in advance, while JIT participants selected content only a moment in advance. Future studies should be designed to clearly separate out the effects of JIT, PS, on-demand, and the duration between content selection and completion. While intervention content was inspired by evidence-based stress reduction strategies and similar to digital micro-interventions tested elsewhere, the particular content had not been tested prior to the present study. Future work should test intervention content and delivery timing separately.

The stress metric employed for the JIT condition was not refined prior to study implementation, and therefore may have prompted interventions at inopportune times. For example, although the stress metric incorporated the number of calendar events per day, it was not capable of distinguishing between work-related and personal events. As personal events could have included self-care activities with stress-reducing impacts, the assumption that a greater number of calendar events per day was associated with greater stress may not have been fully accurate. Additional system limitations included: (1) PS participants were required to manually schedule interventions for themselves without the assistance of a calendar integration; and (2) eight participants had to turn off their cameras due to heavy system load, which constrained the stress metric employed for JIT participants.

Privacy and Ethics

User privacy is a major concern with any application that tracks user behavior. Privacy in the context of work-related stress is even more sensitive, since, in a toxic work environment, stress-related concerns can be stigmatized [129]. Hence, privacy regarding tracking stress-related data is very sensitive, and it must be well regulated within respective organizations. Note that for inferring stress, we used high-level activity data from each participant (e.g., the total number of emails in a given window, the total number of minutes in meetings, etc.). Such data pose relatively few privacy challenges. Irrespective of the granularity of such data, strong regulations need to be established regarding this data collection. In addition, ethical decisions about when to intervene, the granularity of intervention (e.g., individual level, community level, etc.), and how such interventions align with the individual preferences for receiving interventions need to be well thought out through user-centered design and ethical review boards within our respective communities. We intend to embrace all of these challenges in our efforts to confront workday stress and all of its adverse side effects for information workers.

5.2 EVALUATION OF INTERVENTION ENGAGEMENT AT WORK

Our prior study with Huey (Section 5.1) revealed that just-in-time (JIT) micro-interventions are effective in significantly reducing momentary stress. We also found participants in both PS and JIT conditions desired an automated system that would be intelligent about the right timing to send the nudges and the right intervention for the given moment – a feature promised by JITAI systems [292]. Just-in-time adaptive intervention (JITAI) systems have been proposed to optimize precision support and personalized intervention delivery, with a promise to deliver the most effective intervention at the moment that the person is most receptive [292, 355]. Operationalizing the adaptation of the system requires choosing appropriate tailoring variables and intervention options to drive the system's decision-making on when to intervene and what to intervene with. Although the use of passively sensed data for contextual understanding and system adaptation is often recommended for the design of JITAI systems [47], many JITAI systems still rely mostly on ecological momentary assessments (EMAs), app usage, or simple temporal features [431, 48]. Passive sensing technologies, however, can offer numerous sets of contextual variables (e.g., location, calendar, movement, activity), and the challenge in designing sensing-capable JITAI systems lies in identifying a core set of tailoring variables that the system should consider for optimizing effective engagement [460].

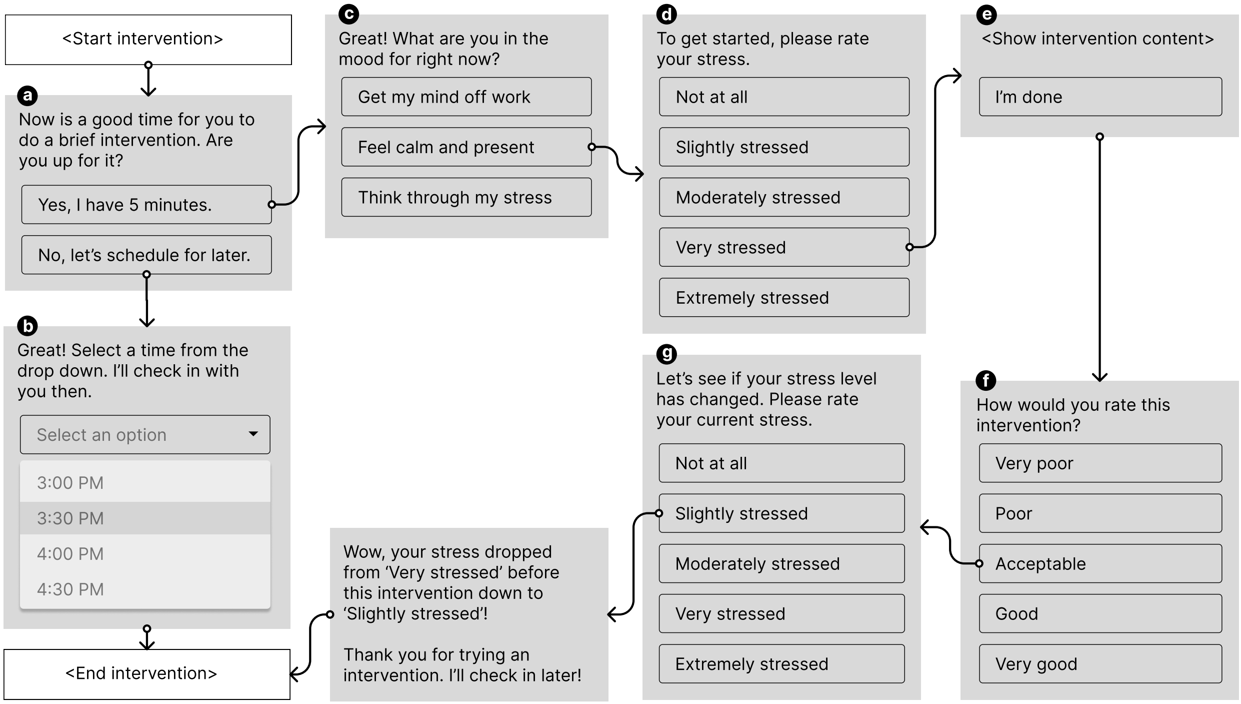
Prior research has examined factors for engagement in digital mental health interventions (DMHIs) [122, 53], such as personal (e.g., demographics, personalities), content-related (e.g., perceived fit, usefulness), and technology-related factors (e.g., technical issues, privacy). However, most of the works reviewed investigate study-long engagement rather than *in-the-moment* engagement factors. Momentary engagement factors have been increasingly important for improving the usability of interventions in real-world contexts [36], especially since the integration of intervention usage into life is a core facilitator for engagement [53]. Despite the promises of JITAI systems to provide precision support [47], very few studies demonstrate the value of just-in-time support in improving user engagement [170, 48]. Therefore, our goal and key contribution is to identify key tailoring variables that influence *momentary* engagement in digital interventions to inform the design of JITAI systems.

Unlike laboratory experiments or controlled studies, our study allows observing users' moment-by-moment interactions with the system in naturalistic work environments. We leverage passively sensed data to gain a glimpse into participants' daily work activities (e.g., emails, meetings, computer activity) and how the work context influences their engagement patterns. In addition, the system usage and telemetry, including which intervention participants chose and liked, can be leveraged to adapt the behavior of the system. Such contextual and intervention data, along with individual demographics, can reveal information about the appropriate conditions that lead to momentary intervention engagement and positive outcomes.

To understand factors that influence the engagement and efficacy of workplace stress-reduction JITAI systems, we analyzed data from the four-week deployment study with Huey, a just-in-time (JIT)[4] micro-intervention system described in Section 5.1.1. Drawing inspiration from prior study findings, we identified five research questions to guide the analysis (Section 5.2.1). We combined participant characteristics data from 43 participants with system usage and telemetry data to contextualize 1,585 system-initiated interventions (Section 5.2.2). From statistical modeling of the impact of individual, contextual, and intervention-related factors on engagement outcomes (Section 5.2.2), we confirmed that individual factors (e.g., age, gender, personality

---

[4] This system did not adapt the intervention content based on changing state or user context and, therefore, is not a full JITAI system.



Now is a good time for you to do a brief intervention. Are you up for it?

Great! Select a time from the drop down. I'll check in with you then.

Great! What are you in the mood for right now?

To get started, please rate your stress.

How would you rate this intervention?

Let's see if your stress level has changed. Please rate your current stress.

Wow, your stress dropped from 'Very stressed' before this intervention down to 'Slightly stressed'!

Thank you for trying an intervention. I'll check in later!

Figure 5.9: System-initiated intervention flow. (a) System sends a nudge to users to perform an intervention. (b) Users can opt to postpone the intervention at a later time. (c) If users choose to do an intervention, they can select from one of three intervention categories. (d) Users first self-report their current stress level. (e) The system shows the intervention content for users to interact with. (f) User rates the intervention. (g) Users self-report their stress levels after the intervention.

traits, coping skills), as well as contextual and content-related factors (e.g., availability, intervention modality), significantly influenced momentary intervention engagement, intervention choice, user ratings, and stress reduction outcomes (Section 5.2.3). These findings suggest tailoring guidelines for JITAI systems whereby contextual and personalized factors can be used to find a positive balance between user preferences and maximal intervention efficacy (Section 5.2.4).

5.2.1 Motivating the Research Questions

There are multiple considerations that influence JITAI systems' decision points (i.e., "a time at which an intervention decision is made" [292]), such as the right timing for a prompt, the right intervention for the moment, or the intervention likely to be engaged in. In the deployment study, decisions are not made entirely by the JITAI systems. For example, users take several turns with the system, and each turn-taking requires an explicit input (i.e., decision) from the user (Figure 5.9). Therefore, tailoring JITAI systems requires not only considerations of contexts but also the impact of user decisions on subsequent interactions. In this section, we draw inspiration from the deployment study and prior findings from Section 5.1.2 to motivate five research questions surrounding the engagement and efficacy of system-initiated interventions.

In Section 5.1.2, we found that participants desired intelligent timing based on their availability and task context. To characterize the timing of system nudges that would increase the likelihood of engagement, we first aimed to understand what contextual factors (e.g., workload) and individual characteristics (e.g., personality) lead to engagement in the system-initiated interventions (RQ1).

RQ1. What factors are associated with the engagement in system-initiated interventions?

The system design allowed participants to choose an intervention category when the bot prompted them to do an intervention (Figure 5.9c). Here, we found that participants had differing opinions about their intervention preferences which might influence their choices. As a follow-up, we explored how different conditions might influence their choices for intervention categories (RQ2).

RQ2. What factors are associated with the choice of intervention categories?

Once a category was chosen, the system randomly selected the least used intervention for the participant to engage in. The interventions could be one of three modalities, and some interventions could be done at the desk while others suggested moving about indoors or outdoors. These characteristics of interventions, such as location, modality, and effort level, are important to consider because certain interventions might not be feasible in certain situations (e.g., the participant cannot go outside) [364, 87]. To understand the consequences of the intervention choice, both the category chosen by the participant and the specific intervention randomly chosen by the system, we aimed to identify contextual, individual, and intervention-related factors on the engagement of that particular intervention (RQ3).

RQ3. Once chosen, what factors are associated with the engagement in the chosen interventions?

Finally, we aimed to understand how contextual, individual, and intervention-related factors affected the user's rating of the interventions (RQ4) and the effectiveness of interventions on stress reduction (RQ5).

RQ4. Once engaged, what factors are associated with the rating of interventions?

RQ5. Once engaged, what factors are associated with the (momentary) effectiveness of interventions?

5.2.2 Data, Processing, and Analysis

To understand which factors contribute to participants engaging in a system-initiated intervention and the effectiveness of interventions at a given moment, we incorporated several data sources. We leveraged per-participant demographics and validated scale responses (coping skills and personality traits) to explain individual differences (Participant Characteristics). From participants' interactions with the system, we determined whether a system-initiated intervention was completed and the effectiveness and rating of interventions once engaged (System Interactions). We harnessed passively sensed data streams to explain the context surrounding the system initiation and intervention engagement (Passively Sensed Context).

This section describes each data source, the motivation for using a subset of variables within each data source, the definition and processing of extracted metrics, and the evaluation of the metrics that led to the final set of variables used for modeling. We then describe our analysis approach. Table 5.1 describes the full set of variables selected for our analysis and their descriptive statistics.

Participant Characteristics

We included demographic variables and personality traits that have been shown to impact engagement [53]. We also included participant coping skills that we assumed to be stable throughout the 4-week study.

Age and gender 67.4% of the 43 participants identified as male. Three participants self-reported to be 18-25 years old, 11 to be 26-35 years old, 18 to be 36-45 years old, eight to be 46-55 years old, two to be 56-65 years old, and one to be 66+ years old. Because of the small sample sizes on either end of the age groups, we combined the lower two and upper three age groups to create more balanced age groups.

Coping skills We included emotion regulation [161] measured by two primary components. Cognitive reappraisal measures a person's ability to change the view of the situation. Expressive suppression measures a person's ability to control the expressive emotional response. Both components range from 1 to 7, with 7 representing high reappraisal and high suppressive ability. We included resilience [379], ranging from 1 to 5 with 5 being high resilience, as the ability to

Table 5.1: List of the per-participant, per-half hour, per-nudge, and per-intervention variables and their descriptive statistics.

	Variable	**Distribution**
Per-participant	*Age*	18-35 (14), 36-45 (18), 46+ (11)
	Gender	Man (29), Woman (14)
	Cognitive Reappraisal	$\bar{x}$=4.69, σ=1.09, min=2, max=7
	Expressive Suppression	$\bar{x}$=3.74, σ=1.26, min=2, max=6
	Resilience	$\bar{x}$=3.51, σ=0.89, min=2, max=5
	Agreeableness	$\bar{x}$=3.79, σ=0.74, min=2, max=5
	Conscientiousness	$\bar{x}$=4.12, σ=0.83, min=2, max=5
	Extraversion	$\bar{x}$=2.62, σ=0.86, min=1, max=4
	Neuroticism	$\bar{x}$=2.85, σ=1.04, min=1, max=5
	Openness	$\bar{x}$=3.48, σ=0.79, min=2, max=5
	Engagement Skewness	$\bar{x}$=-0.09, σ=0.58, min=-1.41, max=1.41
Per-half hour	*Nudge Probability*	$\bar{x}$=0.06, σ=0.04, min=0.02, max=0.30
Per-nudge	*Meeting Counts*	$\bar{x}$=0.30, σ=0.51, min=0, max=3
	No Meeting Minutes	$\bar{x}$=5.43, σ=11.79, min=0, max=50
	Self Event Counts	$\bar{x}$=0.12, σ=0.37, min=0, max=3
	Email Messages Sent	$\bar{x}$=0.39, σ=0.88, min=0, max=8
	Email Messages Read	$\bar{x}$=3.22, σ=5.09, min=0, max=54
	Chat Messages Count	$\bar{x}$=3.88, σ=5.90, min=0, max=64
	Adhoc Call Count	$\bar{x}$=0.04, σ=0.21, min=0, max=2
	Number of Attention Signals	$\bar{x}$=1442.64, σ=1316.68, min=0, max=5705
	Nudge Source	JIT algorithm (1337), Rescheduled (248)
	Engaged	TRUE (563), FALSE (1022)
Per-intervention (chosen)	*Category*	Address (112), Calm (338), Distract (200)
	Engaged	TRUE (563), FALSE (87)
	Modality	Video (113), Prompt (384), Conversation (24)
	Location	At desk (422), Inside (84), Outside (15)
Per-intervention (completed)	*Stress Reduction*	$\bar{x}$=0.29, σ=0.53, min=-1, max=3
	Rating	$\bar{x}$=3.61, σ=1.02, min=1, max=5
	Improved	TRUE (150), FALSE (371)
	Liked	TRUE (289), FALSE (232)
	Stress Before	$\bar{x}$=2.14, σ=0.97, min=1, max=5
	Category	Address (77), Calm (275), Distract (72)
	Modality	Video (60), Prompt (340), Conversation (24)
	Location	At desk (325), Inside (84), Outside (15)

bounce back or recover from stress. Both emotion regulation and resilience measure a person's ability to cope with stressors.

Personality traits We also included the Big Five personality traits [99, 272, 205] because they are known to impact stress [97, 52] and engagement in mental health treatment [53]. Agreeableness is described as being friendly, caring, likable, and altruistic. Conscientiousness is related to being competent, organized, thorough, and deliberate. Extraversion is related to being sociable, energetic, adventurous, and enthusiastic. Neuroticism is described as being anxious, irritable, self-conscious, and vulnerable. Openness is associated with being curious, imaginative, unconventional, and interested in a variety of things. Openness, neuroticism, and agreeableness are generally associated with higher engagement in DMHIs [53]. Each personality trait ranges from 1 to 5, with 5 representing the high presence of the trait.

System Interactions

From system usage data, we reconstructed each participant's step-by-step interaction with the system as represented in Figure 5.9.

Engagement label We labeled each system-initiated intervention as "engaged" in an intervention (i.e., *Engaged* = TRUE) if the participant explicitly marked the intervention as done (i.e., clicking on "I'm done" button in Figure 5.9e), regardless of whether they completed any subsequent prompts (i.e., Figure 5.9f-g). Any ignored, incomplete, or timed-out nudges were considered not engaged.

As shown in Figure 5.9b, participants can opt to postpone the intervention to a later time. Each system-initiated intervention that was triggered at this later requested time was categorized as "rescheduled" (i.e., *Nudge Source* = Rescheduled). Although these rescheduled nudges looked identical to JIT nudges, we hypothesized that the participants would be more likely to engage in a system-initiated intervention if they remembered having postponed the intervention to a time that is more suitable for engagement.

We found that 91.3% of nudges were responded to within 30 minutes of the nudge.

Each nudge is canceled by the system if there had not been a participant response in 30 minutes and if there were outstanding EMA requests. Once participants began conversing with the bot, 96.8% of the engaged interventions were completed within 30 minutes of the first interaction with the bot. To better capture the context surrounding the intervention engagement, we associated each engagement label with a timestamp from the first interaction with the bot. For cases where the participant ignored the nudge, we associated each engagement label with a timestamp of the initial nudge.

Because our goal is to understand the influence of workplace contextual factors, we limited our analysis to weekdays (Monday-Friday). We further excluded days that participants explicitly stated as being out of the office for the entire day to remove noise introduced by atypical weekdays. As a result, there was a total of 1,585 system-initiated interventions with available context data, and 35.5% (563) of those nudges led to a completed intervention engagement.

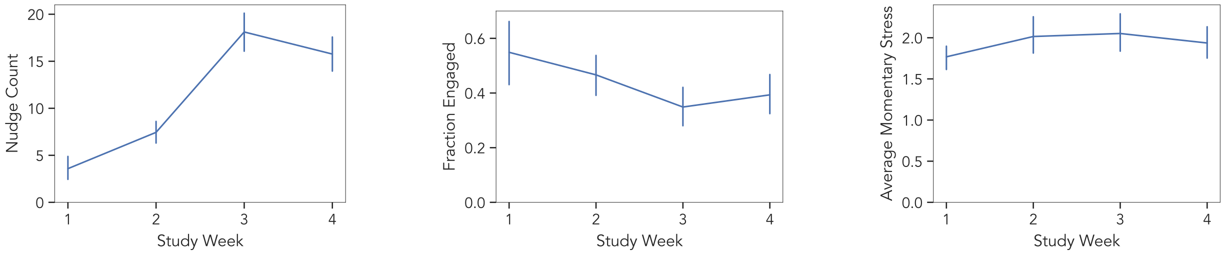
Intervention choice When the system sends a nudge to perform an intervention, users can choose from three intervention categories (Figure 5.9c). For simplicity, we refer to these three intervention categories as 'Distract' (i.e., 'Get my mind off work'), 'Calm' (i.e., 'Feel calm and present'), and 'Address' (i.e., 'Think through my stress'), respectively, in the rest of the paper. There were 18 interventions per category.

If a participant chose an intervention category (Figure 5.9c), we marked that nudge with binary labels of *Distract Chosen*, *Calm Chosen*, or *Address Chosen*. Each of these interventions was further labeled with *Modality* and *Location* based on the specific intervention that the system chose within the category. *Modality* is one of three options – video, prompt, or conversation. Of all the interventions presented to participants, there were 113 video-based, 384 prompt-based, and 24 conversation-based interventions. *Location* refers to whether the intervention could be performed at the desk (e.g., watching a nature video), anywhere inside (e.g., running cold water on your hands), or outside (e.g., taking a walk). There were 422 interventions that were done at the desk, 84 inside, and 15 outside. There were 650 total instances with category labels, and 563 (86.6%) of those were “engaged” interventions.

Engagement outcome Of 563 system-initiated interventions that were engaged, 521 (92.5%) had both pre- and post-intervention stress ratings and intervention ratings. 275 of these interventions engaged were Calm interventions, 169 were Distract interventions, and 77 were Address interventions.

We looked at two outcome metrics – momentary *Stress Reduction* and intervention *Rating*. The average stress rating before engaging in the intervention was 2.14 (σ=0.97) on a 5-point scale (1=Not at all stressed; 5=Extremely stressed). On average, each intervention engagement led to a momentary stress reduction of 0.28 (σ=0.53) and a rating of 3.61 (σ=1.02; 1=Very poor; 5=Very good). We further binarized these outcome metrics to determine if a certain intervention engagement *Improved* stress (i.e., self-reported stress rating was lowered after the intervention use) and if the participant *Liked* the intervention (i.e., rated as ‘Good’ or ‘Very good’). In our data, *Stress Reduction* and *Rating* were positively correlated (Pearson r=0.214, p$\ll$0.001). As such, the binarized outcomes, *Improved* and *Liked*, were also significantly associated (χ^2(1)=18.839, p$\ll$0.001).

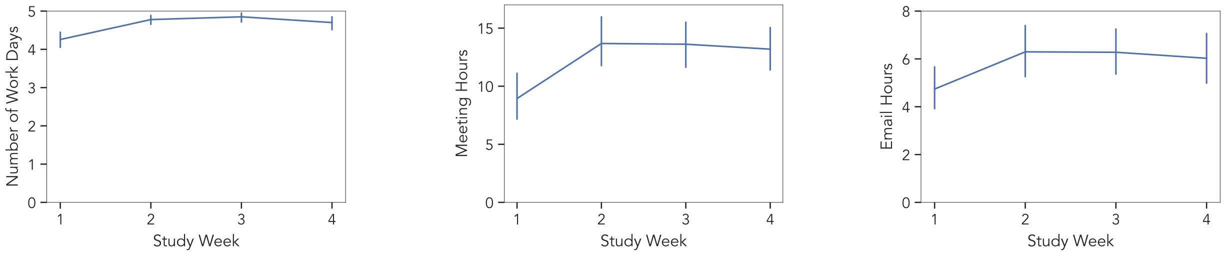
System-initiated intervention probability Throughout the 4 week study, the average number of interventions in which participants engaged increased from 1.61 (σ=1.66) per participant at week 1 to 5.84 (σ=3.98) at week 4. This increase could be plausibly attributed to the increased number of nudges sent to the participants as the study progressed (Figure 5.10a). Week 1 of the study happened to be the week following a major US holiday (4th of July) when many employees were more likely



(a) Number of system-initiated interventions

(b) Fraction of intervention nudges engaged

(c) Average momentary stress rating



(d) Number of work days

(e) Number of meeting hours

(f) Number of email hours

Figure 5.10: Weekly trend of (a) the number of system-initiated interventions, (b) the number of intervention engagements, (c) the average momentary stress ratings (1=Not at all; 2=Slightly stressed; 3=Moderately stressed), (d) the number of work days and workload, (e) the number of hours spent in meetings, and (f) the number of hours spent in attending to emails. Error bars indicate 95% confidence intervals from bootstrapping (n=1000).

to recharge after vacations [444]. The average number of work days per participant during week 1 was 4.33 (σ=1.02), compared to 4.90 during week 3 (σ=0.37, Figure 5.10d). Our data revealed that the general workload was lower immediately following the holidays. In fact, the meeting volume significantly increased throughout the study (β=1.269, F(1)=7.479, p=0.007, Figure 5.10e). The average total hours of meetings per week increased from 8.94 (σ=6.41) at week 1 to 13.19 (σ=6.38) at week 4 (t(82.0)=-3.048, p=0.003, Figure 5.10e). The average total hours of email per week increased from 4.74 (σ=2.91) at week 1 to 6.03 (σ=3.52) at week 4 (t(79.2)=-1.827, p=0.071, Figure 5.10f).

The system determined the threshold for triggering interventions based on the individual baselines captured at week 1. Both the increased workload and self-reported stress ratings are likely to have caused the average number of nudges per participant to increase after the first week. Although the number of intervention engagements increased with the increased number of nudges, the fraction of nudges that led to engagement decreased over time with 0.549 (σ=0.385) at week 1 to 0.393 (σ=0.242) at week 4 (Figure 5.10b), indicating that there may be an upper limit to the number of nudges that will lead to engagements. In fact, the weekly number of nudges and the fraction of engagements were negatively correlated (Pearson r=-0.319, p=1.786e-06). Therefore, when modeling the factors that contribute to engagement in system-initiated interventions, we also consider the general probability of receiving a nudge in the following way.

On average, participants received 48.37 (σ=15.29) system nudges and performed 18.2 (σ=9.48) interventions during the 4-week study. Although each of these nudges could be considered as an independent, repeated observation, system-initiated interventions were not uniformly distributed throughout the day due to particulars of the system design. The system nudges partially depended on self-reported stress levels which were captured through EMAs. These EMAs were spread out evenly between the fixed start and end work hours of each participant, typically 8 AM and 5 PM. Because the system's JIT algorithm runs every 5 minutes to check if a nudge needs to be sent based on the stress score and the EMA stress ratings, the most likely hours for receiving a system nudge is shortly after the EMA, leading to each participant receiving more nudges during certain hours of the day than others (Figure 5.11a). To account for such variability in receiving system-initiated interventions, we incorporate the momentary nudge probability in our analysis. Because the range of nudge probabilities is small ($\bar{x}$=0.06, σ=0.04), we multiply the measure by 100 to represent it in percentages.

(a) Probability of receiving a nudge

(b) Intervention engagement density

Figure 5.11: (a) Probability of receiving a system-initiated intervention during typical work hours (8 AM - 5 PM). Error bars indicate 95% confidence intervals from bootstrapping (n=1000). (b) Kernel density estimate for engaging in a system-initiated intervention. Blue lines indicate participants for whom the Fisher-Pearson coefficient of skewness is significant ($p < 0.05$).

Temporal engagement skewness Prior research has found that different hours of the day were seen as good or not-so-good timing for stress interventions [355]. In Section 5.1.2, we found that participants typically scheduled interventions at the beginning or the end of the day, indicating that there may be individual preferences for when to engage in interventions. To examine if a certain participant has a temporal tendency to engage, we computed the Fisher-Pearson coefficient of skewness, or *Engagement Skewness*, on the hourly intervention engagements per participant. A positive *Engagement Skewness* means that the participants tend to engage at the beginning of their workday, and a negative *Engagement Skewness* means that the participants tend to engage towards the end of their workday. We used the skewness metric instead of simply looking at the engagement during the morning and the afternoon to account for individual differences in working hours. We found that 61.90% of the participants skewed towards engaging at the end of the day while 38.10% of the participants skewed towards the beginning of the day. Figure 5.11b shows the probability density distribution of intervention engagement per participant. We incorporated this skewness per participant in our analysis.

Passively Sensed Context

To understand the context surrounding a system-initiated intervention, we leveraged two sources of passively sensed data that required little or no action from the participants to collect.

One source is the custom sensing software that ran on participants' desktops to infer stress scores. The sensing software also captured general user computer activity events such as mouse and keyboard interactions into a single metric, *Number of Attention Signals*, which could be an important indicator for presence. We hypothesized that presence at the computer could lead to higher engagement in interventions as the nudges were designed to grab the attention of participants at work. The range of values for *Number of Attention Signals* was fairly large (max=5705) compared to other variables, so we divided the values by 2000 to estimate a comparable coefficient and confidence intervals during modeling (i.e., to have odds ratios within 2 decimal points). When interpreting the effect sizes, we corrected for this factor of 2000. We also hypothesized that the likelihood of engaging in a stress-reduction intervention during active participation in a meeting is low. Therefore, we included *No Meeting Minutes* to represent the total number of minutes without a scheduled meeting with others and *Self Event Count* as the total number of calendar events with only the participant as the attendee.

The other is data collected via Microsoft's Viva Insights[5]. As part of the enrollment process for the study, we obtained explicit consent from all participants to join their study data with their Viva Insights data. Viva Insights captures de-identified activity aggregates in 30-minute windows for Microsoft tools across all devices associated with an individual's work account. From this source, we included *Meeting Counts*, *Adhoc Call Count*, and *Email Messages Sent/Read*. We excluded *Chat Messages Count* from

---

[5] https://docs.microsoft.com/en-us/workplace-analytics/use/metric-definitions

our analysis because the nudges were delivered through Teams, and our data source cannot be used to discern if the messages were coming from the bot. Comparison of means revealed that there was a significant difference in *Chat Messages Count* between when participants were engaged and when they were not engaged in an intervention ($t(1312.6){=}{-}3.452$, p$\ll$0.001).

Because our data is limited to half-hourly windows, we associated the contextual metrics with each system-initiated intervention by taking the half-hour window that holds the nudge timestamp.

Analysis

All research questions (Section 5.2.1) are aimed at estimating the effects of contextual, individual, and intervention characteristics on binary outcome variables (e.g., *Engaged*, *Distract Chosen*, *Liked*). Thus, we built a logistic regression model predicting each outcome based on a combination of per-participant characteristics, per-half-hour nudge probability, per-nudge contextual metrics, or per-intervention metrics as fixed effects. The outputs of the logistical regression models are presented as odds ratios in Table 5.2-Table 5.6, representing a ratio of odds (e.g., probability of engaging vs probability of not engaging) under two different conditions (e.g., being a woman vs not being a woman). The data processing was conducted using Python packages (e.g., numpy, pandas, scipy, seaborn) and the models were tested using R libraries (e.g., lme4, car, performance).

We determined the significance of the fitted model against the null hypothesis model using the analysis of deviance. We conducted analyses of variance to estimate the significance of fixed effects. Because the data is unbalanced (i.e., unequal number of observations for each level of a factor), we obtain Type II sums of squares [230]. For categorical variables of 3 levels or more (e.g., *Age Group*, *Category*) that were found

Table 5.2: RQ1: Odds ratios and confidence intervals for each predictor of *Engaged* for all system-initiated nudges. (* p<0.05, ** 0.05<p<0.01, *** 0.01<p<0.001).

	Engaged	
Predictors	**Odds Ratios**	**CI**
(Intercept)	**0.10*****	0.03 – 0.30
Age Group (36-45)	1.17	0.88 – 1.56
Age Group (46+)	**1.47***	1.07 – 2.02
Gender (Woman)	**1.37***	1.05 – 1.79
Cognitive Reappraisal	**1.14***	1.01 – 1.29
Expressive Suppression	0.92	0.83 – 1.02
Resilience	1.05	0.88 – 1.25
Agreeableness	0.87	0.72 – 1.05
Conscientiousness	1.13	0.96 – 1.33
Extraversion	1.03	0.89 – 1.21
Neuroticism	0.97	0.83 – 1.12
Openness	1.15	0.99 – 1.34
Engagement Skewness	**0.64*****	0.51 – 0.79
Nudge Probability	1.02	1.00 – 1.05
Meeting Counts	**0.62*****	0.49 – 0.78
No Meeting Minutes	1.00	0.99 – 1.01
Self Event Counts	1.15	0.85 – 1.54
Email Messages Sent	1.05	0.92 – 1.20
Email Messages Read	1.01	0.99 – 1.03
Adhoc Call Count	0.89	0.52 – 1.49
Number of Attention Signals	**1.39*****	1.17 – 1.66
Trigger Source (Rescheduled)	**1.77***	1.32 – 2.38
Observations = 1585; Tjur's R^2=0.073		

to be significant, we estimated pairwise differences using Tukey's HSD procedure. Multicollinearity in fixed effects was tested using the variance of inflation factor (VIF), and none of our models exhibited a multicollinearity issue. All reported effects are statistically significant (p < 0.05) unless noted otherwise.

5.2.3 Results

RQ1: Effects on intervention engagement

To answer RQ1, we modeled *Engaged* as a function of per-participant characteristics, per-half hour nudge probability, and per-nudge contextual metrics. The logistic regression model of *Engaged* with all fixed effects was significantly different from the null hypothesis model (i.e., *Engage* $\sim$ 1; $\chi^2(21)$=115.5, p$\ll$0.001). Table 5.2 outlines the odds ratios and confidence intervals for each predictor. We found *Gender*, *Cognitive Reappraisal*, *Engagement Skewness*, *Meeting Count*, *Number of Attention Signals*, and *Nudge Source* to have statistically significant effects according to Type II analyses of variance ($p < 0.05$).

Reviewing the coefficients of our fixed effects, we found that being 46+ years old, being a woman, higher *Cognitive Reappraisal*, higher *Number of Attention Signals*, and receiving a rescheduled nudge led to a higher likelihood of intervention engagement. For predictors that increased the likelihood of engagement, receiving a rescheduled nudge has the largest effect size, with a 77% increase in the likelihood of engaging in the intervention compared to JIT algorithm-based nudges. Being in the 46+ age group has a 47% increase in the likelihood of engaging compared to being in the baseline (i.e., 18-35 age group). An increase of 2000 attention signals has a 39% increase in the likelihood of engaging. Being a woman has a 37% increase in the likelihood of engaging than not being a woman (or being a man, since our sample only included two gender types). An increase of one point in cognitive reappraisal skills would result in a 14% increase in the likelihood of engaging.

On the other hand, higher *Engagement Skewness* and *Meeting Count* led to a lower likelihood of intervention engagement. For predictors that decreased the likelihood of engagement, an increase of one unit in *Engagement Skewness* would lead to a 36% decrease (i.e., $1 - 0.64 = 0.36$) in the likelihood of engaging. In plain terms, the more the participant tends to engage at the beginning of the day, the less likely they would engage in interventions. Having one more meeting has a 38% decrease in the likelihood of engaging. Inversely, having one less meeting leads to a 61% increase (i.e., 1 / 0.62 = 1.61) in the likelihood of engaging.

RQ2: Effects on intervention choice

To answer RQ2, we modeled each of the three outcome measures – *Distract Chosen*, *Calm Chosen*, and *Address Chosen* – as a function of per-participant characteristics, per-half hour nudge probability, and per-nudge contextual metrics. Table 5.3 outlines the odds ratios and confidence intervals for all models.

Likelihood of choosing Distract interventions

The logistic regression model of *Distract Chosen* was found to be significant ($\chi^2(21)$=75.65, p≪0.001). Reviewing the coefficients of our fixed effects, we found that being 36-45 years old increased the likelihood of choosing Distract interventions by 65%. Being 36-45 years old had a higher likelihood of Distract interventions than being 46+ years old and being 18-35 years old, but only the differences from being 46+ years old were statistically significant. An increase of one point in *Openness* scale had a 67% increase in the likelihood of choosing Distract interventions. On the other hand, being a woman and an increase of 2000 attention signals decreased the likelihood of choosing Distract interventions by 38% and 35%, respectively.

Likelihood of choosing Calm interventions

The logistic regression model of *Calm Chosen* was found to be significant ($\chi^2(21)$=38.641, p=0.011). We found that an increase of one point in cognitive reappraisal skills has a 27% increase in the likelihood of choosing Calm interventions. On the other hand, an increase of one point in *Openness* has a 29% decrease in the likelihood of choosing Calm interventions.

Table 5.3: RQ2: Odds ratios and p-values for each predictor of *Distract Chosen*, *Calm Chosen*, *Address Chosen* for all participant chosen interventions (* p<0.05, ** 0.05<p<0.01, *** 0.01<p<0.001).

	Distract Chosen		**Calm Chosen**		**Address Chosen**	
Predictors	*Odds Ratios*	*CI*	*Odds Ratios*	*CI*	*Odds Ratios*	*CI*
(Intercept)	**0.11***	0.01 – 0.73	1.84	0.32 – 10.62	0.44	0.04 – 4.52
Age Group (36-45)	1.65*	1.03 – 2.67	0.62*	0.40 – 0.96	0.88	0.46 – 1.65
Age Group (46+)	0.71	0.40 – 1.26	0.95	0.57 – 1.57	1.6	0.85 – 3.07
Gender (Woman)	**0.62***	0.39 – 0.98	1.31	0.87 – 1.96	1.29	0.73 – 2.26
Cognitive Reappraisal	1.01	0.82 – 1.25	**1.27***	1.06 – 1.53	**0.67***	0.53 – 0.84
Expressive Suppression	0.99	0.83 – 1.19	1.06	0.90 – 1.25	0.95	0.75 – 1.18
Resilience	0.93	0.69 – 1.25	0.87	0.67 – 1.13	**1.42***	1.00 – 2.02
Agreeableness	0.75	0.55 – 1.03	1.02	0.77 – 1.36	1.46	0.99 – 2.20
Conscientiousness	1.30	0.98 – 1.75	0.92	0.71 – 1.18	0.74	0.51 – 1.05
Extraversion	1.18	0.91 – 1.53	1.00	0.79 – 1.27	0.79	0.57 – 1.11
Neuroticism	0.88	0.68 – 1.14	1.00	0.79 – 1.26	1.36	0.99 – 1.89
Openness	**1.67****	1.30 – 2.16	**0.71***	0.57 – 0.88	0.83	0.60 – 1.13
Engagement Skewness	1.12	0.74 – 1.68	1.11	0.78 – 1.59	0.86	0.53 – 1.38
Nudge Probability	1.00	0.96 – 1.05	1.01	0.97 – 1.05	0.97	0.92 – 1.03
Meeting Counts	0.84	0.55 – 1.25	1.07	0.76 – 1.52	1.14	0.71 – 1.77
No Meeting Minutes	1.00	0.99 – 1.02	1.00	0.99 – 1.01	1.00	0.98 – 1.02
Self Event Counts	1.08	0.63 – 1.79	0.83	0.53 – 1.30	1.19	0.65 – 2.05
Email Messages Sent	1.05	0.84 – 1.30	1.00	0.82 – 1.22	0.99	0.74 – 1.30
Email Messages Read	1.03	0.99 – 1.07	0.99	0.96 – 1.03	0.96	0.90 – 1.01
Adhoc Call Count	0.88	0.36 – 1.94	0.93	0.45 – 1.93	1.26	0.45 – 3.05
Number of Attention Signals	**0.65****	0.48 – 0.87	1.18	0.92 – 1.53	1.33	0.94 – 1.87
Trigger Source (Rescheduled)	1.01	0.61 – 1.63	0.76	0.49 – 1.16	1.39	0.81 – 2.35
Observations = 650		Tjur's R^2=0.115		Tjur's R^2=0.059		Tjur's R^2=0.089

Likelihood of choosing Address interventions

The logistic regression model of *Address Chosen* was found to be significant ($\chi^2(21)$=53.352, p$\ll$0.001). We found that an increase of one point in *Resilience* increases the likelihood of choosing Address interventions by 42%, while an increase of one point in cognitive reappraisal skills decreases the likelihood of Address interventions by 33%.

Table 5.4: RQ3: Odds ratios, confidence intervals, and p-values for each predictor of *Engaged* for all interventions after participants chose a category (* p<0.05, ** 0.05<p<0.01, *** 0.01<p<0.001).

	Engaged after chosen	
Predictors	**Odds Ratios**	**CI**
(Intercept)	**0.05***	0.00 – 0.94
Age Group (36-45)	1.06	0.54 – 2.05
Age Group (46+)	1.65	0.71 – 4.02
Gender (Woman)	1.65	0.88 – 3.16
Cognitive Reappraisal	1.11	0.83 – 1.48
Expressive Suppression	1.25	0.95 – 1.57
Resilience	**1.63***	1.05 – 2.59
Agreeableness	1.24	0.77 – 2.00
Conscientiousness	0.81	0.55 – 1.17
Extraversion	0.72	0.48 – 1.08
Neuroticism	1.39	0.96 – 2.07
Openness	1.24	0.88 – 1.76
Engagement Skewness	1.01	0.60 – 1.73
Nudge Probability	1.01	0.95 – 1.07
Meeting Counts	0.65	0.41 – 1.06
No Meeting Minutes	1.00	0.98 – 1.02
Self Event Counts	1.94	0.86 – 5.61
Email Messages Sent	1.16	0.84 – 1.67
Email Messages Read	0.97	0.93 – 1.03
Adhoc Call Count	0.54	0.23 – 1.33
Number of Attention Signals	0.90	0.61 – 1.33
Trigger Source [Rescheduled]	1.79	0.85 – 4.10
Category [Calm]	1.14	0.60 – 2.14
Category [Address]	0.55	0.22 – 1.43
Modality [prompt]	**3.53***	1.35 – 9.60
modality [video]	**5.86***	1.69 – 21.68
Location [Inside]	**0.43***	0.23 – 0.83
Location [Outside]	0.34	0.11 – 1.18
	Observations = 650; Tjur's R^2=0.12	

RQ3: Effects of intervention choice on engagement

To answer RQ3, we modeled *Engaged* as a function of per-participant characteristics, per-half hour nudge probability, per-nudge contextual metrics, and per-intervention characteristics. The logistic regression model of *Engaged* was significant ($\chi^2(27)$=69.595, p$\ll$0.001, Table 5.4) We found that an increase of one point of *Resilience* would increase the likelihood of engaging by 63%. The effects of *Modality* were more pronounced. Having a prompt-based and video-based intervention increased the likelihood of engaging by 253% and 486%, respectively, compared to having a conversation-based intervention. Pairwise comparison for *Modality* also confirmed that conversation-type interventions were less likely to lead to engagement compared to video or prompt-based interventions. Having an intervention that could be performed inside, but not at the desk, decreased the likelihood of engaging by 57%. Inversely, having an intervention that could be performed at the desk increased the likelihood of engaging by 133% (i.e., $1/0.43 = 2.33$) compared to one that could be done indoors.

RQ4: Effects on intervention rating

To answer RQ4, we modeled *Liked* as a function of per-participant characteristics, per-half hour nudge probability, per-nudge contextual metrics, and per-intervention metrics. Because we wanted to account for the potential impact of the intervention effects on its rating, we included *Stress Reduction* as another fixed effect. The logistic regression model of *Liked* was significant ($\chi^2(28)$=92.385, p$\ll$0.001, Table 5.5).

We found that being a woman, higher *Cognitive Reappraisal*, higher *Extraversion*, and higher *Stress Reduction* led to a higher likelihood of liking the intervention. We found pronounced effects of *Gender* and *Stress Reduction*. Being a woman or having a reduction in stress by one point results in the increased likelihood of liking the intervention by 151% and 136%, respectively. An increase of one point in *Extraversion* or an increase of one point in *Cognitive Reappraisal* increases the likelihood of liking the intervention by 46% and 35%, respectively.

On the other hand, being 36-45 years old, higher *Expressive Suppression* and higher *Nudge Probability* led to participants being less likely to like the intervention. Being 36-45 years old decreases the likelihood of liking the intervention by 49% compared to being 18-35 years old. An increase of one point in *Expressive Suppression* decreases

Table 5.5: RQ4: Odds ratios, confidence intervals, and p-values for each predictor of *Liked* for all interventions that participants engaged in (* p<0.05, ** 0.05<p<0.01, *** 0.01<p<0.001).

	Liked	
Predictors	**Odds Ratios**	**CI**
(Intercept)	5.95	0.51 – 70.91
Age Group [36-45]	**0.51***	0.30 – 0.86
Age Group [46+]	1.01	0.55 – 1.83
Gender [Woman]	**2.51*****	1.51 – 4.20
Cognitive Reappraisal	**1.35****	1.08 – 1.69
Expressive Suppression	**0.82***	0.68 – 0.99
Resilience	1.10	0.80 – 1.51
Agreeableness	0.72	0.50 – 1.02
Conscientiousness	0.80	0.58 – 1.09
Extraversion	**1.46****	1.10 – 1.95
Neuroticism	0.84	0.64 – 1.11
Openness	0.99	0.75 – 1.30
Engagement Skewness	0.73	0.47 – 1.12
Nudge Probability	**0.94****	0.90 – 0.99
Meeting Counts	0.93	0.61 – 1.42
No Meeting Minutes	1.00	0.99 – 1.02
Self Event Counts	0.66	0.40 – 1.10
Email Messages Sent	1.01	0.79 – 1.28
Email Messages Read	1.00	0.96 – 1.04
Adhoc Call Count	0.66	0.24 – 1.69
Number of Attention Signals	1.16	0.84 – 1.59
Trigger Source [Rescheduled]	0.78	0.47 – 1.30
Category [Calm]	0.80	0.49 – 1.29
Category [Address]	0.94	0.44 – 2.01
Stress Reduction	**2.36*****	1.60 – 3.54
Modality [prompt]	0.52	0.17 – 1.53
modality [video]	0.75	0.22 – 2.47
Location [Inside]	0.69	0.40 – 1.19
Location [Outside]	2.56	0.79 – 9.27
Observations - 521; Tjur's R^2=0.166		

the likelihood of liking interventions by 18%. An increase of one standard deviation (i.e., 4 percentage points) in *Nudge Probability* would decrease the likelihood of liking the interventions by 22% (i.e., $e^{4\times\ln(0.94)} = 0.78$)).

Table 5.6: RQ5: Odds ratios, confidence intervals, and p-values for each predictor of *Improved* for all interventions that participants engaged in. (* p<0.05, ** 0.05<p<0.01, *** 0.01<p<0.001).

	Improved	
Predictors	**Odds Ratios**	**CI**
(Intercept)	**0.00*****	0.00 – 0.00
Age Group [36-45]	0.58	0.30 – 1.13
Age Group [46+]	0.98	0.49 – 1.98
Gender [Woman]	**0.41****	0.21 – 0.77
Cognitive Reappraisal	**0.69****	0.52 – 0.91
Expressive Suppression	1.01	0.79 – 1.28
Resilience	0.84	0.55 – 1.25
Agreeableness	**1.73***	1.10 – 2.76
Conscientiousness	1.27	0.86 – 1.89
Extraversion	0.89	0.62 – 1.29
Neuroticism	**0.57****	0.39 – 0.81
Openness	1.09	0.78 – 1.55
Engagement Skewness	1.18	0.68 – 2.03
Nudge Probability	**1.09****	1.03 – 1.16
Meeting Counts	0.85	0.48 – 1.42
No Meeting Minutes	1.01	0.99 – 1.03
Self Event Counts	1.29	0.69 – 2.36
Email Messages Sent	1.04	0.78 – 1.39
Email Messages Read	0.96	0.90 – 1.01
Adhoc Call Count	0.94	0.28 – 2.63
Number of Attention Signals	1.40	0.95 – 2.05
Trigger Source [Rescheduled]	1.08	0.57 – 2.02
Category [Calm]	**0.43****	0.23 – 0.78
Category [Address]	**0.40***	0.16 – 0.97
Stress Before	**5.76****	3.98 – 8.64
Rating	**2.47****	1.84 – 3.39
Modality [prompt]	**6.65***	1.53 – 36.45
modality [video]	**5.62***	1.12 – 34.10
Location [Inside]	1.17	0.60 – 2.25
Location [Outside]	0.23	0.03 – 1.10
	Observations =521; Tjur's R^2=0.338	

RQ5: Effects on intervention effectiveness

To answer RQ5, we modeled *Improved* as a function of per-participant characteristics, per-half hour nudge probability, per-nudge contextual metrics, and per-intervention metrics. We include *Stress Before* to account for the effects of pre-intervention stress rating on the effectiveness of the intervention and *Rating* to account for the potential effect of participants' rating of the intervention on their post-intervention stress rating.

The logistic regression model of *Improved* was significant ($\chi^2(30)$=186.66, p$\ll$0.001, Table 5.5).

We found that higher *Agreeableness*, higher *Nudge Probability*, higher *Stress Before*, higher *Rating*, and getting prompt-based or video interventions led to participants being more likely to improve on their stress rating. The effects of *Modality*, *Stress Before*, and *Rating* were the most pronounced. Having a prompt-based or video-based intervention increases the likelihood of improving by 565% and 462%, respectively, compared to a conversation-based intervention. An increase of one point in the stress rating before doing the intervention increases the likelihood of improving by 476%. An increase of one point in the user rating for the intervention increases the likelihood of improving by 147%. An increase of one point in *Agreeableness* also increases the likelihood of improving by 73%. For *Nudge Probability*, an increase of one standard deviation (i.e., 4 percentage points) would increase the likelihood of improving by 41% (i.e., $e^{4\times\ln(1.09)} = 1.41$).

On the other hand, being a woman, higher *Cognitive Reappraisal*, higher *Neuroticism*, and choosing Calm or Address interventions led to participants being less likely to improve their stress rating. Being a woman decreases the likelihood of improving by 59% compared to being a man. Choosing a Calm or Address intervention decreases the likelihood of improving by 57% and 60%, respectively, compared to choosing a Distract intervention. Inversely, choosing a Distract intervention increases the likelihood of improving by 133% (i.e., $1/0.43 = 2.33$) and 150% (i.e., $1/0.4 = 2.5$) compared to Calm and Address interventions, respectively. An increase of one point in *Cognitive Reappraisal* and an increase of one point in *Neuroticism* decreases the likelihood of improvement by 31% and 43%, respectively.

5.2.4 Discussion

In this study, we leveraged the survey, EMA, and passively-sensed data from the deployment of the JIT stress-reduction intervention system (Figure 5.9) to identify significant individual, contextual, and intervention-related factors that influence the momentary engagement, the choice of interventions, the engagement given an intervention choice, the user rating of interventions engaged, and the stress reduction from the engagement. In summarizing our findings, we first categorize these factors into (1) non-modifiable individual factors, (2) modifiable individual factors, (3) contextual factors, and (4) content factors.

The distinction between modifiable and non-modifiable individual factors is important for intervention design. Once modifiable factors are identified, strategies can be deployed to directly influence those factors if those strategies can lead to a greater impact on the engagement or the efficacy of the interventions. Non-modifiable factors are also important to determine which subset of populations can benefit from additional targeted support through organizational or policy-level changes [188]. In our analysis, non-modifiable individual factors include gender and age, and modifiable individual factors, although debatable [337], include emotion regulation skills and personality traits. Our findings revealed that these individual factors not only influence study-long engagement [53] but also influence momentary engagement and can be useful for tailoring JITAI systems. Contextual factors such as workload (e.g., meeting counts, email counts) and availability (e.g., presence or activity at the computer) are helpful in the JITAI system's decision-making process for when to interrupt the user. Content factors include intervention-related information such as the amount of effort required, the modality of intervention delivery, and the ideal location for intervention engagement, and these factors can inform JITAI systems in determining which intervention to present to the user given the understanding of the current context.

In this section, we summarize and discuss the findings organized by these four categories of factors with recommendations for design and future research.

Non-modifiable individual factors

Tailoring JITAI systems to individuals has been suggested to improve engagement and efficacy of interventions in prior research [431]. In our analysis, we found several significant effects of gender and age on our outcome measures. Participants who self-reported as being a woman had over twice the likelihood of liking interventions than being a man, despite having less likelihood of improving from engaging in interventions. Being a woman also had 38% less likelihood of choosing Distract interventions than being a man, suggesting that there may be an unobserved motivational factor. For example, although not statistically significant, women reported higher momentary stress on average compared to men (2.15 for women vs 1.84 for men; $t(20.756)$=-1.545, p=0.137) which may contribute to choosing more Calm or Address interventions to reduce their stress. Although the general findings from DMHI studies that women are more likely to engage in digital interventions than men are corroborated by our analysis [53], the fact that women improve less despite engaging more, liking interventions more, and choosing more Address (i.e., high reward) interventions is a concern for systems design that only take engagement metrics into account.

Participants that are 46+ years old were 47% more likely to engage in interventions while being 36-45 years old increased the likelihood of choosing Distract interventions by 65% and decreased the likelihood of liking interventions by 49% compared to those that are 18-35 years old. Although the effects of different age groups on engagement have mixed results across prior studies, the higher engagement rate for 46+-year-old participants in the study could be explained by a higher rate of interest in digital interventions for older populations [258]. Participants in the 36-45-year-old age group choosing Distract interventions and not liking them highlight an opportunity for finding different types of interventions that they might enjoy.

Recommendation #1: We recommend that intervention content be tailored based on gender and age. At the same time, the efficacy and user rating of interventions should be closely monitored to ensure that certain groups enjoy the same level of benefits as other groups. New intervention content could be co-designed with groups that may not be benefiting as much and added to the system on a regular basis to equalize outcomes across subgroups.

Modifiable individual factors

Our findings corroborate prior research that personality trait is a strong factor in the engagement of DMHIs [53]. Prior study has found that openness to experience is associated with better adherence and lower odds of attrition [279]. In our analysis, we did not find a significant effect on engagement, but we found that *Openness* significantly increases the likelihood of choosing Distract, but decreases the likelihood of choosing Calm interventions. One possible explanation may be the variety in the intervention content, which people with high openness would prefer [271]. Distract interventions tend to offer more variety in content with videos of travel destinations and humor as well as opportunities to explore social connections, whereas Calm interventions were mostly introspective activities such as breathing or focused observations.

Although the effects were only moderately significant ($p < 0.061$), we found that an increase by one point in *Neuroticism* and *Agreeableness* scales increases the likelihood of choosing Address interventions, which were designed to help users directly address and problem-solve stress-inducing components of their lives. For participants with high scores in *Neuroticism* and *Agreeableness* scales, it is possible that interventions that help them directly address their stress were more appealing than others that were designed to distract from stress and refocus on the present. Prior study has also found neuroticism and agreeableness to associate with a stronger interest in the use of stress management apps [133]. We also found that one point increase in *Agreeableness* scale was associated with an increased likelihood of improvement by 73%, whereas the same point increase in *Neuroticism* scale was associated with a lower likelihood of improvement by 43%, despite both having higher tendency to choose Address interventions. Since agreeableness is known to be positively associated with the therapeutic alliance in mental health treatments [67], it is possible that the prosocial and cooperative nature of those with high agreeableness [158] allowed them to fully engage in the Address interventions that were more action-oriented, leading to a greater improvement. On the other hand, neuroticism has been known to negatively correlate with adherence to mental health recommendations [311] and to a wide variety of mental health treatment outcomes [67].

Recommendation #2: We recommend that JITAI systems carefully monitor potentially unhelpful usage behaviors by incorporating personality traits in the adaptation algorithm because they may impact the choice of interventions that may lead to negative downstream effects on outcomes. For example, it may be beneficial to offer a variety of more effortful interventions for people who report higher scores in openness. For people who report higher scores in neuroticism, the system could suggest less effortful interventions or explore different types of interventions that may lead to more stress reduction.

Across the board, emotion regulation had significant effects. One point increase in the *Cognitive Reappraisal* scale was associated with a 14%[6] increase in the likelihood

---

[6] It is important to note the differences in the range of point scales. An increase of 14% for a 7-point scale is equivalent to an increase of 20% for a 5-point scale.

of engaging, a 27% increase in the likelihood of choosing Calm interventions, a 35% increase in the likelihood of liking the intervention, a 33% decrease in the likelihood of choosing Address interventions, and a 31% decrease in the likelihood of improvement after engagement. On the other hand, one point increase in the *Expressive Suppression* scale was associated with an 18% decrease in the likelihood of liking the intervention. We also found that one point increase in the *Resilience* scale increases the likelihood of choosing Address interventions by 42% and increases the likelihood of subsequently engaging in interventions by 63%.

Prior research has explored the role of emotion regulation in stress coping. For example, emotion regulation skills help assess the stressful situation and determine the appropriate emotional response [290] or act as buffers against the negative effects of stress [432]. Emotion regulation has also been theorized as a moderator for increased resilience after encountering a stressful situation [413]. Although our analysis cannot claim the causal direction between coping skills and engagement, our findings suggest that emotion regulation and resilience may play a role in, not only the stress-coping process but in choosing different interventions or deciding to engage in an intervention. The role of emotion regulation strategies in altering our decisions and choices in various contexts has been empirically studied in highly controlled laboratory settings [159]. While most research has argued for increasing coping skills as an outcome measure or a treatment target [42], our findings suggest promising new research directions in understanding how coping skills could impact our decisions to engage in therapeutic interventions in the moment.

Recommendation #3: We recommend further research to explore how taking a dynamic approach to personality and coping skills would inform the design of JITAI systems. Prior research has studied personality traits [337, 154] and coping skills [42] as “states” with intrapersonal variations and targeted their modifications through therapeutic strategies (e.g., CBT). Because of the potential mediating role of coping skills in perceived stress [56] and the role of perceived stress in outcome improvement (i.e., one point increase in stress rating before the intervention leads to being over 5 times more likely to improve in our findings), coping skills should be measured periodically and incorporated into the decision-making process of JITAI systems.

Contextual factors

We found that the more the participant tends to engage at the end of the day, the more likely they would engage in interventions and that the system-initiated interventions that were rescheduled to a later time increased the likelihood of engagement. These findings may suggest a tendency to defer interventions to later in the day. Prior work that applied the Self Determination Theory [108] to JITAI systems suggested that perceived competence and self-regulation abilities may deplete throughout the day, potentially leading to unhealthy choices (e.g., unhealthy food, alcohol) towards the end of the day [386]. This has a serious consequence for those that tend to procrastinate or postpone healthy behaviors (e.g., exercise, stress intervention) towards the end of the day. On the other hand, allowing people to defer an intervention to a specific time may increase self-efficacy and perception of control, which might lead to an increased chance of behavior change [7]. In fact, in our qualitative data, we found that many participants liked the ability to perform the interventions when they wanted. We also found that the increase in *Nudge Probability* decreases the likelihood of liking the intervention but increases the likelihood of improvement, revealing that a JIT intervention might be “a bitter pill to swallow” but a useful pill nonetheless. However, a relentless reminder could lead to distraction [227] and eventual system abandonment [48].

Recommendation #4: Future JITAI systems should carefully balance individual preferences with intervention efficacy and help users discover what works best for them. Our findings show encouragement that intelligent intervention timing based on contextual information could improve the engagement and effectiveness of interventions, but perhaps at the cost of users liking the interventions or leading to unhealthy choices towards the end of the day. It should also consider that higher user ratings may not always reflect the effectiveness of the interventions and aim to explore both the rating and the improvement in determining the timing of interventions.

In evaluating the work context, as expected, we found that a nudge sent at a time when the user is less likely to be in a meeting but more active at the desk improves engagement. Contrary to our hypothesis, we found no significant associations with *No Meeting Minutes* or *Self Event Counts*. It is possible that there is high variability in the level of focus and attention needed during times carved out for self. For example, our data sources cannot discern if the times carved out for self were work-related (e.g., focus time for reading, writing) or nonwork-related (e.g., running errands, child pick up, exercise). The current study's data sources cannot achieve automatic detection of activities beyond basic work activities, such as meetings, emails, chats, calls, or computer activities. Although tailoring to the activity context is the defining promise of JITAI systems, automatically detecting the activities performed within a time window is not an easy task. In addition, the impact of accuracy in detection on engagement and outcomes is unknown.

Recommendation #5: Further research is necessary to understand the impact of accurate activity detection and intelligent timing on the engagement and effectiveness of JITAI systems.

Content factors

In our analysis, intervention categories, modalities, and locations showed pronounced effects on engagement and improvement, suggesting the importance of the intervention content in the design of JITAI systems. We found that having an intervention that could be done at the desk more than doubled the likelihood of engagement compared to an intervention that could be done indoors, but not at the desk. It is possible that leaving the desk at the moment of the nudge was not appropriate given the situation or there was an unobserved motivational barrier. In these scenarios, suggesting a different activity, rescheduling the activity or waiting until the next appropriate transition time might have been beneficial. To improve engagement, intervention designers could provide additional desk-based stress reduction techniques to minimize the burden of leaving the desk.

On the other hand, we found that having a prompt- or video-based intervention increased the likelihood of engagement by over 3-fold and improvement by over 5-fold, compared to a conversation-based intervention. Although prompt- and video-based interventions were typically less effortful than conversation-based interventions that require many turn-taking interactions with the bot, conversation-based interventions were designed to address the sources of the stress with the hope of creating a longer-lasting impact. It is possible that the conversations were not usable for participants to fully engage in the content. This finding suggests that quick, effortless interventions could be useful at the moment, but complex, turn-taking interventions need more thoughtful redesign.

Recommendation #6: We recommend that JITAI systems provide a variety of interventions to fine-tune its recommendations based on contexts but also to identify interventions that may need redesign.

Limitations

Our analysis setup does not allow for determining the causal relationship between the individual, contextual, and content factors with engagement, stress reduction, and intervention ratings. Even though the deployment study collected long-term stress measures via the DASS-21 scale, our analysis was limited to momentary stress ratings because micro-interventions are more appropriate for proximal outcomes over distal outcomes [36]. Micro-randomized trials (MRTs) are a promising research direction for JITAI systems to quantify the impact of tailored interventions on both short- and long-term outcomes [217] with careful considerations for the appropriate sample size [240]. Our data was also limited by a small sample population that exhibited low-stress levels. Prior work has suggested that, when evaluating engagement (or attrition), the severity of symptoms should also be considered [122]. Therefore, future research should evaluate the system through MRTs with a sample population exhibiting high severity of stress symptoms within. Finally, workplace stress-reduction JITAI systems and their deployment must be evaluated along with the unobservable or difficult-to-modify external factors (e.g., workplace culture [123], power asymmetry and surveillance [196]) that could have a non-negligible impact on the engagement and the efficacy of these systems.

5.3 SUMMARY OF CONTRIBUTIONS TO THESIS

Reducing workplace stress is of critical importance for employees and employers alike, but workplace stress-reduction interventions have produced mixed results due to engagement and adherence barriers. To improve the engagement of interventions at work, this chapter describes how we made adaptations to existing EBPIs and directly integrated them into the context of work as just-in-time micro-interventions.

We designed and developed a workplace stress-reduction intervention system that uses passively sensed information to deliver these interventions. Then we conducted a four-week longitudinal study testing the delivery timing and content type of these micro-interventions with information workers and analyzed the data from just-in-time uses of the system to identify individual, contextual, and intervention-related factors that influence engagement and efficacy of the interventions.

We found that digital micro-interventions are effective at reducing momentary workplace stress. In analyzing the usage data, we found that personality traits and coping skills influenced momentary engagement, the choice of interventions, the user rating, and the efficacy of interventions. We also found that sending intervention nudges at inferred moments of high stress was associated with liking the interventions less but increased the likelihood of stress reduction. The intervention content types had the strongest effects with prompt-based and video-based interventions being associated with a higher likelihood of improvement by over five times that of the conversation-based interventions.

Our findings suggested that our micro-interventions should be integrated into workplaces now for immediate, positive impact and that tailoring of delivery timing based on momentary contextual information, personalizing intervention contents, and balancing user and system control may improve user engagement and stress reduction outcomes. The findings also highlighted several design and research opportunities for JITAI systems. We recommend that JITAI systems provide a variety of intervention content types and tailor the content and delivery based on non-modifiable individual factors (e.g., gender, age) as well as potentially modifiable individual factors (e.g., personality "states" and coping skills). We also see opportunities for further research in understanding the impact of personality and coping skills on JITAI system usage and the potentially dynamic nature of such factors in the context of JITAI systems. We recommend that the design of JITAI systems employ a multi-faceted approach that incorporates individual, contextual, and intervention-related factors into the decision-making process of intervention timing, intervention content, and interpretation of the subjective ratings.

My work in the workplace stress context demonstrated that *the adaptation of EBPIs into digital micro-interventions* combined with the use of passively sensed and momentary contextual information to deliver them at the moment *can improve the engagement of stress-reduction interventions* in the workplace. Furthermore, the usage data can be harnessed to *continuously improve* the engagement in and the effectiveness of JITAI systems by tailoring and adapting its decision-making process to the individual and their context.

Chapter 6

DISCUSSION AND CONCLUSION

6.1 SUMMARY OF THESIS CONTRIBUTIONS

My dissertation research aimed to improve engagement in mental health and well-being care by taking a holistic stance in the examination of the situated contexts surrounding the engagement. I hypothesized that the barriers to engagement are situated in personal and population-level contexts, and these challenges are rooted in tensions among multiple needs observable through human-centered and computational methods. Grounded on the understanding of these needs, I claimed that we can improve engagement by rethinking and adapting existing interventions.

My dissertation demonstrated this approach in three well-being contexts: (1) the COVID-19 pandemic, (2) co-morbid cancer and depression, and (3) workplace stress. In the COVID-19 context, I showed how using computational methods, grounded on holistic frameworks, could reveal changes in the multi-system of human needs and disparities in changes on the US population scale (Chapter 3). In the co-morbid cancer and depression context, I leveraged the parallel journeys framework to reveal tensions between cancer and psychosocial care journeys and developed a collaborative behavioral activation system to reimagine how technology could enhance collaborative depression care experiences (Chapter 4). In the workplace stress context, I translated evidence-based psychosocial interventions into digital micro-interventions that could be directly integrated into the workplace and evaluated how momentary contextual information through passive sensing could power the delivery of the right intervention at the right time (Chapter 5).

My work in these well-being contexts followed three key research activities that are typically seen in the human-centered design process [176, 278]: (1) framing and understanding, (2) designing and adapting, and (3) deploying and evaluating.

First, I examined situated contexts holistically to discover tensions across needs and resources. In order to frame the holistic understanding, I developed a computational framework – human needs framework – to characterize the shifts within the multi-level system of human needs (Section 3.1). I leveraged the social determinants of health to identify the most likely socioeconomic and environmental factors (e.g., median household income) associated with differential resource tensions introduced by the pandemic (Section 3.2). I also developed a conceptual design framework – parallel journals framework – to characterize the challenges that arise when parallel cancer and psychosocial care processes collide (Section 4.1). Using two independent contexts, I demonstrated that the use of computational and human-centered methods combined with holistic frameworks can help identify stakeholder needs and tensions among needs that get in the way of engaging in mental health care and well-being activities.

Second, I identified modification targets to existing interventions that can be enhanced through the use of technology to ease the tensions among needs and to improve engagement. I developed a technology-enhanced collaborative behavioral activation model that facilitates the delivery of core components of behavioral activation and collaborative care for achieving high fidelity while simultaneously supporting patient-provider collaboration, patients' self-efficacy, and continuity of care (Section 4.2). I applied the model in the digitization of behavioral activation and collaborative care and in the iterative design of the SCOPE system. I also translated components of cognitive behavioral therapy and dialectical behavioral therapy into digital stress-reduction micro-interventions that minimize user effort and can fit easily into a busy work schedule (Section 5.1). In addition, these interventions are delivered in the most likely stressful moment based on computational contextual understanding. Using two independent contexts, I demonstrated that contextual understanding can help the appropriate design, adaptation, and delivery of technology-facilitated interventions to reduce barriers and improve engagement.

Lastly, I evaluated my proposed adaptations and technology enhancements to interventions in real-world contexts. The SCOPE system was deployed to pilot participants in two cancer clinics where I found that our system supported all four touch points identified in our model with a potential to bring positive changes to the existing care (Section 4.2). I also found significant social and organizational challenges with the introduction of the system, such as the lack of motivation and justification for using the system and the lack of knowledge around how to incorporate the system into the existing care practices, suggesting the need to further explore additional adjustments to the interventions as well as the implementation of the interventions. The deployment of the workplace stress-reduction intervention system revealed that micro-interventions are effective in momentary stress-reduction (Section 5.1). Further analysis of the engagement data revealed that individual, contextual, and intervention content factors have a significant impact on the rate of engagement, the intervention choices, the user rating after intervention use, and the effectiveness of interventions on stress reduction, highlighting the potential to continuously monitor and tailor the system to maximize engagement and effectiveness. In these two independent contexts, I demonstrated that technology-facilitated interventions must be evaluated in the contexts for which they were designed. The qualitative and quantitative data collected from the deployment of these systems can be harnessed not only to validate the designs but also to inform future adaptations of interventions that can be facilitated by the system (e.g., personalized system recommendations) or by the care team (e.g., process modifications by the BHPs).

These three research activities across three mental health and well-being contexts, therefore, contributed to my thesis:

> *Engagement in mental health and well-being care can be improved through contextual and continuous adaptation of interventions by using human-centered and computational methods to understand stakeholder needs at personal and population scales.*

In summary, my dissertation demonstrated that using human-centered and computational methods can reveal the tensions within the multi-level system of stakeholder and well-being needs that get in the way of engaging in mental health and well-being activities. In my work, such contextual understanding, at both personal and population scales, has led to the appropriate design and adaptation of technology-facilitated interventions necessary to improve engagement in mental health and well-being care. By deploying and evaluating these technology-facilitated interventions in real-world contexts, I have identified future opportunities for adapting the interventions.

6.2 DESIGN RECOMMENDATIONS AND FUTURE DIRECTIONS

My work across three well-being contexts provided design recommendations and future research opportunities relevant to their contexts in each of their respective chapters (Section 3.1.7, Section 3.2.3, Section 4.1.5, Section 4.2.5, Section 5.1.3, Section 5.2.4). To conclude the thesis, I reflect on all of my dissertation work together and present abstract discussion points useful for the design and research of technology-enhanced mental health and well-being interventions.

6.2.1 Consider the holistic context

My commitment to examining the holistic context stems from the unfortunate observation that engaging in mental health and well-being activities is often deprioritized over other activities. In many of the contexts that I studied, I observed that the internal and external pressures from what is typically considered as “basic” needs (e.g., physical health, economic stability) forced people to ignore their mental health needs. Looking at mental health needs alongside other well-being needs highlighted the amount of force that other needs exert on the deprioritization of mental health or on mental health distress and helped us identify where to scaffold support.

For example, during the pandemic, I observed that the shift to basic needs was exponential and that the impact of public policies on social-emotional well-being cannot be ignored. Such observations surfaced because physiological, economic, social, and psychosocial needs were examined together. In addition, by looking at differential engagement behaviors across the SDoH – a holistic approach to viewing health –, I identified intervention design opportunities around specific determinants and vulnerable populations that could benefit from targeted shielding. In cancer settings, I confirmed what we already know about the prioritization of the treatment of the disease process (i.e., cancer) over the holistic needs of the patient as a person. When we brought the psychosocial care journey alongside the cancer care journey, we exposed how patients fall through the cracks at the intersection of the two care journeys and identified cancer-related stressors and treatments as the main source of frequent adjustments to depression care. Therefore, rather than looking at mental health engagement as a problem to solve in isolation, I recommend that mental health be examined alongside other well-being needs.

6.2.2 Design adaptations for multiple temporal granularities

In my work, I studied multiple levels of temporal granularity, from looking at changes in search behaviors across the pandemic, looking at depression care engagement between sessions, to looking at micro-interventions delivered in the moment. Time is an important dimension of analysis for long- and short-term effects of events or interventions. For example, I observed that an earlier shelter-in-place mandate was associated with fewer mental health site visits at first, but a longer shelter-in-place mandate, most likely due to the earlier mandate, was associated with more expressions of negative mental health issues. The target success outcome for the RCT involving SCOPE looks at longitudinal changes in PHQ-9 and GAD-7 scores within patients.

On the other hand, time is an important design material for adapting existing interventions. Typically, evidence-based psychosocial interventions have treatment durations that last many weeks or months with repeated sessions with the BHPs. When designing for patient-provider collaborative behavioral activation interventions, we also aimed to achieve continuity in care. However, our formative study revealed that unexpected circumstances and stressors in cancer settings frequently derail treatment plans and such disruptions are difficult to avoid. In fact, high dropout rates or low adherence to depression treatment are not unique to cancer care settings [303]. Because we understood that patients with cancer may not be seeing BHPs as regularly as needed, we ensured to prioritize digitizing the core components of behavioral activation that patients can use when they are not in sessions. Such adaptation not only fills in the gap during missed appointments or delayed follow-ups but also provides opportunities to engage at a finer temporal granularity (i.e., days instead of weeks).

Furthermore, we can consider situations where there is no follow-up appointment such that the interventions can only be done in a single session. Single-session interventions [359, 358] have gained popularity in recent years to maximize efficacy while intentionally eliminating the need for continuity in treatment. Similarly, our stress-reduction micro-interventions were intentionally designed as one-shot treatments to reduce stress in the moment in the shortest duration possible. Within the context of just-in-time adaptive intervention systems, these micro-interventions are meant to be explored without any sequential order, unlike traditional interventions. Therefore, I recommend time (i.e., duration, frequency, continuity, sequence) as a design dimension to explore in the adaptation and digital translation of evidence-based psychosocial interventions.

6.2.3 Navigate between personalization and standardization

In addition to looking at different temporal granularities, my work also examined contexts at small individual scales, at multistakeholder organizational scales, as well as at large population scales. Large-scale observational analyses are suitable for identifying macro-behavioral patterns that can highlight where to target interventions and monitor the effectiveness of interventions. However, to design population-specific interventions, small-scale focused studies aimed at contextualizing individuals' experiences are necessary [39]. On the other hand, small-scale observational analyses provide the necessary expressivity to describe how multi-level systems of needs interact as well as to present the multi-stakeholder perspectives on the challenges of conflicting needs. At individual levels, interventions can be personalized to fit the needs of one individual without affecting others in the group. However, to measure the clinical efficacy of interventions (i.e., in RCTs), population-level analysis and standardization are often necessary.

Navigating from individual to population scales, therefore, creates an interesting challenge for intervention design. Again, traditional evidence-based psychosocial interventions strive for standardization and linear programming, with fidelity to care as one of the success metrics for implementation. Evaluation metrics are also standardized through clinically validated measures such as PHQ-9 and GAD-7, all of which are used in our RCT. Such standardization is highly useful and efficient for a program or population-level evaluations. However, many of the classifications and evaluation metrics have been criticized for being culturally insensitive [340] and being biased [310]. On the other hand, standardization seems absolutely necessary to provide some type of quality assurance to the explosion of digital mental health technologies we face today. Although my dissertation encourages contextual and continuous adaptation of interventions to meet people where they are, further research is recommended to consider how we should rethink and adapt the evaluation process along the dimensions of individual to population scales.

6.2.4 Understand the role of technology through deployment in real-world contexts

In my examination of the interplay between digital and offline exclusion for digital disparities, of the integration of SCOPE in the collaborative care of patients with cancer and depression, and of the passively sensed data providing just-in-time stress relief, I observed how technology has a profound impact on people's lives and how people's lives have a profound impact on how they use the technology. Such observations were made possible because the technologies were deployed and used in real-world contexts. Especially in the cancer context, I saw how the introduction of technology can disrupt interpersonal dynamics and clinical workflows. At the same time, it has the potential to enhance patient-provider collaboration through asynchronous communication and data sharing. During the pandemic, behavioral data from search engines revealed how technology could be utilized to meet offline demands, but we are also reminded of how exclusion from digital tools could exacerbate offline disparities. One of our motivations for just-in-time micro-interventions is to minimize potential conflicts that may arise for taking personal time at work [123, 184], but over-indexing on technical solutions will not help with promoting psychologically safe work cultures. My dissertation work led to my humble realization that technology must be examined in the context of the dynamic interactions among individual and environmental factors [63].

Often, these technology solutions are decided and deployed by the organizational leaders with the financial resources to provide them as benefits. As a follow-up work beyond my dissertation, I explore both the positive and negative implications of deploying workplace well-being technologies in the organizational context. For example, the presence of sensing technologies that drive contextual understanding and just-in-time decision-making could be misused or perceived as surveillance mechanisms that could amplify mistrust in the organization. Organizational-level sensed measurements themselves may not align with individual well-being goals and definitions, and such misalignment may have cascading negative consequences that impact team dynamics and organizational culture. In addition to exploring how to mitigate the potential harms of workplace well-being technologies, I examine how to design these solutions as tools of employee empowerment, for example, to leverage personal insights in negotiating organizational-level changes that better support employee well-being.

My dissertation work follows the human-centered design process, which does not necessarily dictate nor disallow looking at the design of technology-enhanced interventions from the sociotechnical perspective. Upon reflecting on my work, I posit that a holistic and multistakeholder perspective on mental health and well-being, rather than a disease-centric view, encouraged the discovery of interpersonal, organizational, and societal challenges in relation to the proposed technology solutions. Because such challenges influence the adoption, engagement, and efficacy of technology-enhanced interventions, I recommend that intervention designs be grounded in the sociotechnical context through deployment and using methods that draw out multistakeholder tensions and socioecological perspectives.

BIBLIOGRAPHY

[1] *Diagnostic and statistical manual of mental disorders: DSM-5*. American Psychiatric Publishing, Washington, DC, fifth edition edition, 2013.

[2] A look at how home care product claim preferences have shifted amid the covid-19 pandemic. https://nielseniq.com/global/en/insights/analysis/2020/a-look-at-how-home-care-product-claim-preferences-have-shifted-amid-the-covid-19-pandemic/ (Accessed 2022-10-28), Apr 2020.

[3] Gregory A. Aarons, Amy E. Green, Lawrence A. Palinkas, Shannon Self-Brown, Daniel J. Whitaker, John R. Lutzker, Jane F. Silovsky, Debra B. Hecht, and Mark J. Chaffin. Dynamic adaptation process to implement an evidence-based child maltreatment intervention. *Implementation Science*, 7(1):1–9, 2012. ISSN 17485908. doi: 10.1186/1748-5908-7-32.

[4] Kibrom A Abay, Kibrom Tafere, and Andinet Woldemichael. Winners and losers from COVID-19: Global evidence from Google Search. *World Bank Policy Research Working Paper*, (9268), 2020.

[5] Nancy E Adler and David H Rehkopf. US disparities in health: descriptions, causes, and mechanisms. *Annu. Rev. Public Health*, 29:235–252, 2008.

[6] Stephen O Agboola, Woong Ju, Aymen Elfiky, Joseph C Kvedar, and Kamal Jethwani. The effect of technology-based interventions on pain, depression, and quality of life in patients with cancer: A systematic review of randomized controlled trials. *Journal of Medical Internet research*, 17(3):e65, 2015. doi: 10.2196/jmir.4009.

[7] Icek Ajzen. The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211, 1991.

[8] Bill Albert and Tom Tullis. *Measuring the user experience: collecting, analyzing, and presenting usability metrics.* Newnes, 2013.

[9] Rajwa Alharthi, Benjamin Guthier, Camille Guertin, and Abdulmotaleb El Saddik. A dataset for psychological human needs detection from social networks. *IEEE Access*, 5:9109–9117, 2017.

[10] William E Allen, Han Altae-Tran, James Briggs, Xin Jin, Glen McGee, Andy Shi, Rumya Raghavan, Mireille Kamariza, Nicole Nova, Albert Pereta, et al. Population-scale longitudinal mapping of COVID-19 symptoms, behaviour and testing. *Nature Human Behaviour*, 4(9):972–982, 2020.

[11] Tim Althoff, Kevin Clark, and Jure Leskovec. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics*, 4:463–476, 2016.

[12] Tim Althoff, Ryen W White, and Eric Horvitz. Influence of Pokémon Go on physical activity: study and implications. *Journal of Medical Internet Research*, 18(12):e315, 2016.

[13] Tim Althoff, Eric Horvitz, Ryen W White, and Jamie Zeitzer. Harnessing the web for population-scale physiological sensing: A case study of sleep and performance. In *Proceedings of the 26th international conference on World Wide Web*, pages 113–122, 2017.

[14] Tim Althoff, Rok Sosič, Jennifer L Hicks, Abby C King, Scott L Delp, and Jure Leskovec. Large-scale physical activity data reveal worldwide activity inequality. *Nature*, 547(7663):336–339, 2017.

[15] Tim Althoff, Eric Horvitz, and Ryen W White. Psychomotor function measured via online activity predicts motor vehicle fatality risk. *NPJ digital medicine*, 1 (1):1–2, 2018.

[16] Tim Althoff, Hamed Nilforoshan, Jenna Hua, and Jure Leskovec. Large-scale diet tracking data reveal disparate associations between food environment and diet. *Nature Communications*, 13(1):1–12, 2022.

[17] American Psychiatric Association. Dissemination of integrated care within adult primary care settings: The collaborative care model, 2016.

[18] Nazanin Andalibi, Pinar Oxturk, and Andrea Forte. Sensitive self-disclosures, responses, and social support on instagram: The case of #depression. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing*, 2017. doi: 10.1145/2998181.2998243.

[19] Barbara L Andersen, Robert J DeRubeis, Barry S Berman, Jessie Gruman, Victoria L Champion, Mary Jane Massie, Jimmie C Holland, Ann H Partridge, Kate Bak, Mark R Somerfield, et al. Screening, assessment, and care of anxiety and depressive symptoms in adults with cancer: An american society of clinical oncology guideline adaptation. *Journal of Clinical Oncology*, 32(15):1605, 2014. doi: 10.1200/jco.2013.52.4611.

[20] Monica Anderson and Emily A Vogels. Americans turn to technology during COVID-19 outbreak, say an outage would be a problem. *Pew Research Center*, 2020.

[21] Janine Archer, Peter Bower, Simon Gilbody, Karina Lovell, David Richards, Linda Gask, Chris Dickens, and Peter Coventry. Collaborative care for depression and anxiety problems. *Cochrane Database of Systematic Reviews*, (10), 2012. doi: 10.1002/14651858.cd006525.pub2.

[22] Peter C Austin. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3):399–424, 2011.

[23] ChaeWon Baek, Peter B McCrory, Todd Messer, and Preston Mui. Unemployment effects of stay-at-home orders: Evidence from high-frequency claims data. *The Review of Economics and Statistics*, 103(5):979–993, 2021.

[24] Scott R Baker, Robert A Farrokhnia, Steffen Meyer, Michaela Pagel, and Constantine Yannelis. How does household spending respond to an epidemic? Consumption during the 2020 COVID-19 pandemic. *The Review of Asset Pricing Studies*, 10(4):834–862, 2020.

[25] Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. Designing mobile health technology for bipolar disorder: A field trial of the monarca system. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2627–2636. ACM, 2013. doi: 10.1145/2470654.2481364.

[26] Manuel Barrera, Felipe G. Castro, Lisa A. Strycker, and Deborah J. Toobert. Cultural adaptations of behavioral health interventions: a progress report. *Journal of consulting and clinical psychology*, 81(2):196–205, 2013. ISSN 19392117. doi: 10.1037/a0027085.

[27] Manuel Barrera, Cady Berkel, and Felipe González Castro. Directions for the Advancement of Culturally Adapted Preventive Interventions: Local Adaptations, Engagement, and Sustainability. *Prevention Science*, 18(6):640–648, 2017. ISSN 13894986. doi: 10.1007/s11121-016-0705-9.

[28] Amy M Bauer, Stephen M Thielke, Wayne Katon, Jürgen Unützer, and Patricia Areán. Aligning health information technologies with effective service delivery

models to improve chronic disease care. *Preventive Medicine*, 66:167–172, 2014. doi: 10.1016/j.ypmed.2014.06.017.

[29] Amy M Bauer, Sarah Hodsdon, Suzanne Hunter, Youlim Choi, Jared Bechtel, and John C Fortney. Lessons from the deployment of the spirit app to support collaborative care for rural patients with complex psychiatric conditions. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 772–780. ACM, 2017. doi: 10.1145/3123024.3125610.

[30] Amy M Bauer, Sarah Hodsdon, Jared M Bechtel, and John C Fortney. Applying the principles for digital development: Case study of a smartphone app to support collaborative care for rural patients with posttraumatic stress disorder or bipolar disorder. *Journal of Medical Internet Research*, 20(6):e10048, 2018. doi: 10.2196/10048.

[31] Amy M Bauer, Matthew Iles-Shih, Reza Hosseini Ghomi, Tessa Rue, Tess Grover, Naomi Kincler, Monica Miller, and Wayne J Katon. Acceptability of mhealth augmentation of collaborative care: A mixed methods pilot study. *General Hospital Psychiatry*, 51:22–29, 2018. doi: 10.1016/j.genhosppsych.2017.11.010.

[32] Amy M Bauer, Mark D Williams, Anna Ratzliff, and Jürgen Unützer. Best practices for systematic case review in collaborative care. *Psychiatric Services*, pages appi–ps, 2019. doi: 10.1176/appi.ps.20190008.

[33] Amy M Bauer, Matthew Jakupcak, Matt Hawrilenko, Jared Bechtel, Rob Arao, and John C Fortney. Outcomes of a health informatics technology-supported behavioral activation training for care managers in a collaborative care program. *Families, Systems, & Health*, 39(1):89, 2021.

[34] Fran Baum, Lareen Newman, and Katherine Biedrzycki. Vicious cycles: digital technologies and determinants of health in Australia. *Health promotion international*, 29(2):349–360, 2012.

[35] Amit Baumel, Theresa Fleming, and Stephen M Schueller. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. *Journal of medical Internet research*, 22(10):e20631, 2020.

[36] Amit Baumel, Theresa Fleming, and Stephen M Schueller. Digital micro interventions for behavioral and mental health gains: core components and conceptualization of digital micro intervention care. *Journal of medical Internet research*, 22(10):e20631, 2020.

[37] Elizabeth A Bayliss, John F Steiner, Douglas H Fernald, Lori A Crane, and Deborah S Main. Descriptions of barriers to self-care by persons with comorbid chronic diseases. *The Annals of Family Medicine*, 1(1):15–21, 2003. doi: 10.1370/afm.4.

[38] Roy Bear, Peter W Choate, Gabrielle Lindstrom, et al. Reconsidering maslow and the hierarchy of needs from a first nations' perspective. *Aotearoa New Zealand Social Work Review*, 34(2), 2022.

[39] Elisabeth Beaunoyer, Sophie Dupéré, and Matthieu J Guitton. COVID-19 and digital inequalities: Reciprocal impacts and mitigation strategies. *Computers in human behavior*, 111:106424, 2020.

[40] Dror Ben-Zeev, Emily A Scherer, Rui Wang, Haiyi Xie, and Andrew T Campbell. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. *Psychiatric Rehabilitation Journal*, 38(4):313–313, 2015. doi: 10.1037/prj0000130.

[41] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300, 1995.

[42] Matthias Berking, Peggilee Wupperman, Alexander Reichardt, Tanja Pejic, Alexandra Dippel, and Hansjörg Znoj. Emotion-regulation skills as a treatment target in psychotherapy. *Behaviour research and therapy*, 46(11):1230–1237, 2008.

[43] Donna L Berry, Brent A Blumenstein, Barbara Halpenny, Seth Wolpin, Jesse R Fann, Mary Austin-Seymour, Nigel Bush, Bryant T Karras, William B Lober, and Ruth McCorkle. Enhancing patient-provider communication with the electronic self-report assessment for cancer: A randomized trial. *Journal of Clinical Oncology*, 29(8):1029, 2011. doi: 10.1200/jco.2010.30.3909.

[44] Donna L Berry, Fangxin Hong, Barbara Halpenny, Ann H Partridge, Jesse R Fann, Seth Wolpin, William B Lober, Nigel E Bush, Upendra Parvathaneni, Anthony L Back, et al. Electronic self-report assessment for cancer and self-care support: Results of a multicenter randomized trial. *Journal of Clinical Oncology*, 32(3):199, 2014. doi: 10.1200/jco.2013.48.6662.

[45] Hugh Beyer and Karen Holtzblatt. *Contextual Design: Defining Customer-Centered Systems*, volume 1. Morgan kaufmann, 1998.

[46] Prachi Bhardwaj. Most job search requirements for unemployment benefits are waived. so why do state websites say otherwise? https://money.com/unemployment-benefits-job-search-requirements-coronavirus/ (Accessed 2021-08-17), May 2020.

[47] N Bidargaddi, G Schrader, P Klasnja, J Licinio, and S Murphy. Designing

m-health interventions for precision mental health support. *Translational psychiatry*, 10(1):1–8, 2020.

[48] Niranjan Bidargaddi, Daniel Almirall, Susan Murphy, Inbal Nahum-Shani, Michael Kovalcik, Timothy Pituch, Haitham Maaieh, Victor Strecher, et al. To prompt or not to prompt? a microrandomized trial of time-varying push notifications to increase proximal engagement with a mobile health app. *JMIR mHealth and uHealth*, 6(11):e10123, 2018.

[49] Melanie Birks, Ysanne Chapman, and Karen Francis. Memoing in qualitative research: Probing data and processes. *Journal of Research in Nursing*, 13(1): 68–75, 2008. doi: 10.1177/1744987107081254.

[50] Kirsten Boehner, Janet Vertesi, Phoebe Sengers, and Paul Dourish. How hci interprets the probes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1077–1086, 2007.

[51] Megan Bogia, Chloé Nurik, Andrea Ngan, Bennett Kuhn, Ila Kumar, and Jessa Lingel. Institutional shadow bodies in mental health care information seeking. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 269–272. ACM, 2018. doi: 10.1145/3272973.3274072.

[52] Niall Bolger and Adam Zuckerman. A framework for studying personality in the stress process. *Journal of personality and social psychology*, 69(5):890, 1995.

[53] Judith Borghouts, Elizabeth Eikey, Gloria Mark, Cinthia De Leon, Stephen M Schueller, Margaret Schneider, Nicole Stadnick, Kai Zheng, Dana Mukamel, Dara H Sorkin, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. *Journal of medical Internet research*, 23(3):e24387, 2021.

[54] Taxiarchis Botsis, Gunnar Hartvigsen, Fei Chen, and Chunhua Weng. Secondary use of ehr: Data quality issues and informatics opportunities. *Summit on Translational Bioinformatics*, 2010:1, 2010. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3041534/.

[55] Edwin Boudreaux, Cris Mandry, and Phillip J Brantley. Stress, job satisfaction, coping, and psychological distress among emergency medical technicians. *Prehospital and Disaster Medicine*, 12(4):9–16, 1997.

[56] Chloe C Boyle, Annette L Stanton, Patricia A Ganz, Catherine M Crespi, and Julienne E Bower. Improvements in emotion regulation following mindfulness meditation: Effects on depressive symptoms and perceived stress in younger breast cancer survivors. *Journal of consulting and clinical psychology*, 85(4):397, 2017.

[57] Virginia Braun and Victoria Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006. doi: 10.1191/1478088706qp063oa.

[58] Paula A Braveman, Catherine Cubbin, Susan Egerter, Sekai Chideya, Kristen S Marchi, Marilyn Metzler, and Samuel Posner. Socioeconomic status in health research: one size does not fit all. *Jama*, 294(22):2879–2888, 2005.

[59] Todd Bridgman, Stephen Cummings, and John Ballard. Who built maslow's pyramid? a history of the creation of management studies' most famous symbol and its implications for management education. *Academy of management learning & education*, 18(1):81–98, 2019.

[60] Rob B Briner. Improving stress assessment: Toward an evidence-based approach to organizational stress interventions. *Journal of Psychosomatic Research*, 43(1):61–71, 1997.

[61] Andrei Broder. A taxonomy of web search. In *ACM Sigir forum*, volume 36, pages 3–10. ACM New York, NY, USA, 2002.

[62] Elizabeth Brondolo, Kahaema Byer, Peter J. Gianaros, Cindy Liu, Aric A. Prather, Kamala Thomas, and Cheryl L. Woods-Giscombé. Apa working group on stress and health disparities.(2017). *Stress and Health Disparities: Contexts, Mechanisms, and Interventions among Racial/Ethnic Minority and Low-Socioeconomic Status Populations. Retrieved from http://www.apa.org/pi/health-disparities/resources/stress-report.aspx*, 2017.

[63] Urie Bronfenbrenner. *The ecology of human development: Experiments by nature and design*. Harvard university press, 1979.

[64] Rodger E Broomé. Review of transformation beyond greed: Native self-actualization. 2017.

[65] Marius Brülhart, Valentin Klotzbücher, Rafael Lalive, and Stephanie K Reich. Mental health concerns during the covid-19 pandemic as revealed by helpline calls. *Nature*, 600(7887):121–126, 2021.

[66] Melissa G Bublitz, Natalie Czarkowski, Jonathan Hansen, Laura A Peracchio, and Sherrie Tussler. Pandemic reveals vulnerabilities in food access: confronting hunger amidst a crisis. *Journal of Public Policy & Marketing*, 40(1):105–107, 2021.

[67] Meredith A Bucher, Takakuni Suzuki, and Douglas B Samuel. A meta-analytic review of personality traits and their associations with mental health treatment outcomes. *Clinical Psychology Review*, 70:51–63, 2019.

[68] Lindsey Rose Bullinger, Jillian B Carr, and Analisa Packham. COVID-19 and Crime: Effects of Stay-at-Home Orders on Domestic Violence. Technical report, National Bureau of Economic Research, 2020.

[69] George Bulman and Robert W Fairlie. Technology and education: Computers, software, and the internet. In *Handbook of the Economics of Education*, volume 5, pages 239–280. Elsevier, 2016.

[70] Sebastian Burchert, Mohammed Salem Alkneme, Martha Bird, Kenneth Carswell, Pim Cuijpers, Pernille Hansen, Eva Heim, Melissa Harper Shehadeh, Marit Sijbrandij, Edith Van't Hof, and Christine Knaevelsrud. User-centered app adaptation of a low-intensity e-mental health intervention for Syrian refugees. *Frontiers in Psychiatry*, 10(JAN), 2019. ISSN 16640640. doi: 10.3389/fpsyt.2018.00663.

[71] Stephanie L Burcusa and William G Iacono. Risk for recurrence in depression. *Clinical Psychology Review*, 27(8):959–985, 2007. doi: 10.1016/j.cpr.2007.02.005.

[72] Nigel E Bush, Steven K Dobscha, Rosa Crumpton, Lauren M Denneson, Julia E Hoffman, Aysha Crain, Risa Cromer, and Julie T Kinn. A virtual hope box smartphone app as an accessory to therapy: Proof-of-concept in a clinical sample of veterans. *Suicide & Life-Threatening Behavior*, 45(1):1–9, 2015. doi: 10.1111/sltb.12103.

[73] Andrew C Butler, Jason E Chapman, Evan M Forman, and Aaron T Beck. The empirical status of cognitive-behavioral therapy: a review of meta-analyses. *Clinical psychology review*, 26(1):17–31, 2006.

[74] Daniel Calderon Gomez. The third digital divide and bourdieu: Bidirectional conversion of economic, cultural, and social capital to (and from) digital capital among young people in madrid. *New Media & Society*, page 1461444820933252, 2020.

[75] Andrew M. Campbell. An increasing risk of family violence during the covid-19 pandemic: Strengthening community collaborations to save lives. *Forensic Science International: Reports*, 2020.

[76] Colleen E Carney, Daniel J Buysse, Sonia Ancoli-Israel, Jack D Edinger, Andrew D Krystal, Kenneth L Lichstein, and Charles M Morin. The consensus sleep diary: standardizing prospective sleep self-monitoring. *Sleep*, 35(2):287–302, 2012.

[77] Daniel Carrión, Elena Colicino, Nicolo Foppa Pedretti, Kodi B Arfer, Johnathan Rush, Nicholas DeFelice, and Allan C Just. Neighborhood-level disparities and subway utilization during the covid-19 pandemic in new york city. *Nature communications*, 12(1):1–10, 2021.

[78] Hilde Castelijns, Vera Eijsbroek, A Th Cees, Harm WJ van Marwijk, Christina M van der Feltz-Cornelis, et al. Illness burden and physical outcomes associated with collaborative care in patients with comorbid depressive disorder in chronic medical conditions: A systematic review and meta-analysis. *General Hospital Psychiatry*, 50:1–14, 2018. doi: 10.1016/j.genhosppsych.2017.08.003.

[79] Loredana Cerbara, Giulia Ciancimino, Massimo Crescimbene, Federica La Longa, Maria Rita Parsi, Antonio Tintori, and Rossella Palomba. A nation-wide survey on emotional and psychological impacts of covid-19 social distancing. *European review for medical and pharmacological sciences*, 2020.

[80] Hung-Hao Chang and Chad D Meyerhoefer. Covid-19 and the demand for online food shopping services: Empirical evidence from Taiwan. *American Journal of Agricultural Economics*, 103(2):448–465, 2021.

[81] Serina Chang, Emma Pierson, Pang Wei Koh, Jaline Gerardin, Beth Redbird, David Grusky, and Jure Leskovec. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840):82–87, 2021.

[82] Mary Charlton, Jennifer Schlichting, Catherine Chioreso, Marcia Ward, and Praveen Vikas. Challenges of rural cancer care in the united states. *Oncology*, 29

(9), 2015. URL https://www.cancernetwork.com/oncology-journal/challenges-rural-cancer-care-united-states/page/0/1.

[83] Annie T. Chen, Shuyang Wu, Kathryn N. Tomasino, Emily G. Lattie, and David C. Mohr. A multi-faceted approach to characterizing user behavior and experience in a digital mental health intervention. *Journal of Biomedical Informatics*, 94 (October 2018):103187, 2019. ISSN 15320464. doi: 10.1016/j.jbi.2019.103187. URL https://doi.org/10.1016/j.jbi.2019.103187.

[84] Chih-Lin Chi, W Nick Street, Jennifer G Robinson, and Matthew A Crawford. Individualized patient-centered lifestyle recommendations: An expert system for communicating patient specific cardiovascular risk information and prioritizing lifestyle options. *Journal of Biomedical Informatics*, 45(6):1164–1174, 2012. doi: 10.1016/j.jbi.2012.07.011.

[85] Eun Kyoung Choe, Nicole B Lee, Bongshin Lee, Wanda Pratt, and Julie A Kientz. Understanding quantified-selfers' practices in collecting and exploring personal data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1143–1152. ACM, 2014. doi: 10.1145/2556288.2557372.

[86] Hyunyoung Choi and Hal Varian. Predicting the present with Google Trends. *Economic record*, 88:2–9, 2012.

[87] Woohyeok Choi, Sangkeun Park, Duyeon Kim, Youn-kyung Lim, and Uichin Lee. Multi-stage receptivity model for mobile just-in-time health intervention. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 3(2), jun 2019. doi: 10.1145/3328910. URL https://doi.org/10.1145/3328910.

[88] Merlin Chowkwanyun and Adolph L Reed Jr. Racial health disparities and covid-19—caution and context. *New England Journal of Medicine*, 383(3):201–203, 2020.

[89] Chia-Fang Chung, Kristin Dew, Allison Cole, Jasmine Zia, James Fogarty, Julie A Kientz, and Sean A Munson. Boundary negotiating artifacts in personal informatics: Patient-provider collaboration with patient-generated data. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 770–786. ACM, 2016. doi: 10.1145/2818048.2819926.

[90] Larisa N Cicila, Emily J Georgia, and Brian D Doss. Incorporating internet-based interventions into couple therapy: Available resources and recommended uses. *Australian and New Zealand Journal of Family Therapy*, 35(4):414–430, 2014. doi: 10.1002/anzf.1077.

[91] Olivier Coibion, Yuriy Gorodnichenko, and Michael Weber. Labor markets during the covid-19 crisis: A preliminary view. Technical report, National Bureau of economic research, 2020.

[92] Tim J Cole and Douglas G Altman. Statistics notes: percentage differences, symmetry, and natural logarithms. *Bmj*, 358, 2017.

[93] Rachel A Coleman and Robert A Neimeyer. Assessment of subjective client agency in psychotherapy: A review. *Journal of constructivist psychology*, 28(1): 1–23, 2015.

[94] Thomas W Colligan and Eileen M Higgins. Workplace stress: Etiology and consequences. *Journal of workplace behavioral health*, 21(2):89–97, 2006.

[95] Paul H Coluzzi, Marcia Grant, James H Doroshow, Michelle Rhiner, Betty Ferrell, and Lynne Rivera. Survey of the provision of supportive care services at national cancer institute-designated cancer centers. *Journal of Clinical Oncology*, 13(3): 756–764, 1995. doi: 10.1200/jco.1995.13.3.756.

[96] Comscore. Comscore Explicit Core Search Share Report (Desktop Only) July

2021 vs August 2021, Total U.S. - Desktop Home & Work Locations. https://www.comscore.com/Insights/Rankings (Accessed 2021-11-11).

[97] Jennifer K Connor-Smith and Celeste Flachsbart. Relations between personality and coping: a meta-analysis. *Journal of personality and social psychology*, 93(6): 1080, 2007.

[98] Cary L Cooper and Sue Cartwright. An intervention strategy for workplace stress. *Journal of psychosomatic research*, 43(1):7–16, 1997.

[99] Paul T Costa and Robert R McCrae. *Revised NEO Personality Inventory (NEO-PI-R) and Neo Five-Factor inventory (NEO-FFI)*. Psychological Assessment Resources, 1992.

[100] Shelia R Cotten and Sipi S Gupta. Characteristics of online and offline health information seekers and factors that discriminate between them. *Social science & medicine*, 59(9):1795–1806, 2004.

[101] David Coyle, Gavin Doherty, Mark Matthews, and John Sharry. Computers in talk-based mental health interventions. *Interacting with Computers*, 19(4): 545–562, 2007. doi: 10.1016/j.intcom.2007.02.001.

[102] David Coyle, Nicola McGlade, Gavin Doherty, and Gary O'Reilly. Exploratory evaluations of a computer game supporting cognitive behavioural therapy for adolescents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2937–2946. ACM, 2011. doi: 10.1145/1978942.1979378.

[103] Mark É Czeisler, Rashon I Lane, Emiko Petrosky, Joshua F Wiley, Aleta Christensen, Rashid Njai, Matthew D Weaver, Rebecca Robbins, Elise R Facer-Childs, Laura K Barger, et al. Mental health, substance use, and suicidal ideation during the covid-19 pandemic—united states, june 24–30, 2020. *Morbidity and Mortality Weekly Report*, 69(32):1049, 2020.

[104] Francesco D'Amuri and Juri Marcucci. 'Google It!' Forecasting the US Unemployment Rate with A Google Job Search Index. *FEEM Working Paper Series*, 2010.

[105] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. Predicting depression via social media. In *Seventh international AAAI conference on weblogs and social media*, volume 13, pages 1–10, 2013.

[106] Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 2098–2110, 2016. doi: 10.1145/2858036.2858207.

[107] Elsbeth Marieke De Korte, Noortje Wiezer, Joris H. Janssen, Peter Vink, and Wessel Kraaij. Evaluating an mHealth app for health and well-being at work: Mixed-method qualitative study. *JMIR mHealth and uHealth*, 6(3):1–17, 2018. ISSN 22915222. doi: 10.2196/mhealth.6335.

[108] Edward L Deci and Richard M Ryan. *Intrinsic motivation and self-determination in human behavior.* Springer Science & Business Media, 2013.

[109] Mary Jo Deering, Erin Siminerio, and Scott Weinstein. Issue brief: Patient-generated health data and health it. *Office of the National Coordinator for Health Information Technology*, 20, 2013.

[110] Evangelia Demerouti, Arnold B Bakker, Friedhelm Nachreiner, and Wilmar B Schaufeli. The job demands-resources model of burnout. *Journal of Applied psychology*, 86(3):499, 2001.

[111] Evangelia Demerouti, Akihito Shimazu, Arnold B. Bakker, Kyoko Shimada, and Norito Kawakami. Work-self balance: A longitudinal study on the effects of job demands and resources on personal functioning in japanese working parents.

Work and Stress, 27(3):223–243, 2013. ISSN 02678373. doi: 10.1080/02678373.2013.812353.

[112] Teresa Deshields, Tiffany Tibbs, Ming-Yu Fan, and Marie Taylor. Differences in patterns of depression after treatment for breast cancer. *Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer*, 15(5): 398–406, 2006. doi: 10.1002/pon.962.

[113] Teresa Deshields, Brad Zebrack, and Vicki Kennedy. The state of psychosocial services in cancer care in the united states. *Psycho-Oncology*, 22(3):699–703, 2013. doi: 10.1002/pon.3057.

[114] Drew Desilver. Not all unemployed people get unemployment benefits; in some states, very few do. https://www.pewresearch.org/fact-tank/2020/04/24/not-all-unemployed-people-get-unemployment-benefits-in-some-states-very-few-do/ (Accessed 2021-08-17), 2020.

[115] Paul DiMaggio and Bart Bonikowski. Make money surfing the web? The impact of Internet use on the earnings of US workers. *American Sociological Review*, 73 (2):227–250, 2008.

[116] Paul DiMaggio and Filiz Garip. Network effects and social inequality. *Annual review of sociology*, 2012.

[117] Paul DiMaggio, Eszter Hargittai, et al. From the ‘digital divide’ to ‘digital inequality’: Studying Internet use as penetration increases. *Princeton: Center for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University*, 4 (1):4–2, 2001.

[118] Justin B Dimick and Andrew M Ryan. Methods for evaluating changes in health care policy: the difference-in-differences approach. *Jama*, 312(22):2401–2402, 2014.

[119] Sona Dimidjian, Steven D Hollon, Keith S Dobson, Karen B Schmaling, Robert J Kohlenberg, Michael E Addis, Robert Gallop, Joseph B McGlinchey, David K Markley, Jackie K Gollan, et al. Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of consulting and clinical psychology*, 74 (4):658, 2006. doi: 10.1037/0022-006x.74.4.658.

[120] Gavin Doherty, David Coyle, and John Sharry. Engagement with online mental health interventions: An exploratory clinical study of a treatment for depression. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012. doi: 10.1145/2207676.2208602.

[121] Gavin Doherty, David Coyle, and John Sharry. Engagement with online mental health interventions: An exploratory clinical study of a treatment for depression. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012. URL https://doi.org/10.1145/2207676.2208602.

[122] Gavin Doherty, David Coyle, and John Sharry. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1421–1430, 2012.

[123] Maureen F Dollard. *Psychosocial safety climate: A lead indicator of workplace psychological health and engagement and a precursor to intervention success*. Routledge, 2012.

[124] E Ray Dorsey and Eric J Topol. State of telehealth. *New England Journal of Medicine*, 375(2):154–161, 2016.

[125] Benjamin G Druss, Thomas Bornemann, Yvonne W Fry-Johnson, Harriet G McCombs, Robert M Politzer, and George Rust. Trends in mental health and

substance abuse services at the nation's community health centers: 1998–2003. *American Journal of Public Health*, 98(Supplement_1):S126–S131, 2008. doi: 10.2105/ajph.2005.076943.

[126] A Duarte, J Walker, S Walker, G Richardson, C Holm Hansen, P Martin, G Murray, M Sculpher, and M Sharpe. Cost-effectiveness of integrated collaborative care for comorbid major depression in patients with cancer. *Journal of Psychosomatic Research*, 79(6):465–470, 2015. doi: 10.1016/j.jpsychores.2015.10.012.

[127] Susan Dumais, Robin Jeffries, Daniel M Russell, Diane Tang, and Jaime Teevan. Understanding user behavior through log data and analysis. In *Ways of Knowing in HCI*, pages 349–372. Springer, 2014.

[128] Stefan Rennick Egglestone, Sarah Knowles, Gill Toms, Penny Bee, Karina Lovell, and Peter Bower. Health technologies 'in the wild': Experiences of engagement with computerised cbt. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2016. doi: 10.1145/2858036.2858128.

[129] Daniel Eisenberg, Marilyn F Downs, Ezra Golberstein, and Kara Zivin. Stigma and help seeking for mental health among college students. *Medical Care Research and Review*, 66(5):522–541, 2009.

[130] David Ekers, Lisa Webster, Annemieke Van Straten, Pim Cuijpers, David Richards, and Simon Gilbody. Behavioural activation for depression; an update of meta-analysis of effectiveness and sub group analysis. *PloS One*, 9(6):e100100, 2014. doi: 10.1371/journal.pone.0100100.

[131] Rosenberg Ekman. *What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS).* Oxford University Press, USA, 1997.

[132] Per Engzell, Arun Frey, and Mark D Verhagen. Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, 118(17), 2021.

[133] Mari Ervasti, Johanna Kallio, Ilmari Määttänen, Jani Mäntyjärvi, Markus Jokela, et al. Influence of personality and differences in stress processing among finnish students on interest to use a mobile stress management app: survey study. *JMIR mental health*, 6(5):e10039, 2019.

[134] Gunther Eysenbach. The law of attrition. *Journal of medical Internet research*, 7 (1):e402, 2005.

[135] Jesse R. Fann and Jennifer Sexton. Collaborative psychosocial oncology care models. In *Psycho-Oncology*, pages 1–16. Oxford University Press, London, 2015. doi: 10.1093/med/9780199363315.003.0099.

[136] Jesse R Fann, Donna L Berry, Seth Wolpin, Mary Austin-Seymour, Nigel Bush, Barbara Halpenny, William B Lober, and Ruth McCorkle. Depression screening using the patient health questionnaire-9 administered on a touch screen computer. *Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer*, 18(1):14–22, 2009. doi: 10.1002/pon.1368.

[137] Jesse R Fann, Kathleen Ell, and Michael Sharpe. Integrating psychosocial care into cancer services. *Journal of Clinical Oncology*, 30(11):1178–1186, 2012. doi: 10.1200/jco.2011.39.7398.

[138] Jesse R Fann, Fangxin Hong, Barbara Halpenny, Traci M Blonquist, and Donna L Berry. Psychosocial outcomes of an electronic self-report assessment and self-care intervention for patients with cancer: A randomized controlled trial. *Psycho-Oncology*, 26(11):1866–1871, 2017. doi: 10.1002/pon.4250.

[139] Kenneth D Feigenbaum and Rene Anne Smith. Historical narratives: Abraham maslow and blackfoot interpretations. *The Humanistic Psychologist*, 48(3):232, 2020.

[140] Jonathan S Feinstein. The relationship between socioeconomic status and health: a review of the literature. *The Milbank Quarterly*, pages 279–322, 1993.

[141] Nuno Fernandes. Economic effects of coronavirus outbreak (COVID-19) on the world economy. *Available at SSRN 3557504*, 2020.

[142] David Fletcher and Mustafa Sarkar. Psychological resilience: A review and critique of definitions, concepts, and theory. *European psychologist*, 18(1):12, 2013.

[143] Eric Fong, Barry Wellman, Melissa Kew, and Rima Wilkes. Correlates of the digital divide: Individual, household and spatial variation. In *Office of Learning Technologies, Human Resources Development*. Citeseer, 2001.

[144] National Institute for Clinical Excellence and Great Britain. Guidance on the use of computerised cognitive behavioural therapy for anxiety and depression. Technical report, 2002.

[145] John C Fortney, Jeffrey M Pyne, Timothy A Kimbrell, Teresa J Hudson, Dean E Robinson, Ronald Schneider, William M Moore, Paul J Custer, Kathleen M Grubbs, and Paula P Schnurr. Telemedicine-based collaborative care for post-traumatic stress disorder: A randomized clinical trial. *JAMA Psychiatry*, 72(1): 58–67, 2015. doi: 10.1001/jamapsychiatry.2014.1575.

[146] Adam Fourney, Ryen W White, and Eric Horvitz. Exploring time-dependent concerns about pregnancy and childbirth from search logs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 737–746, 2015.

[147] George H Gallup. Human needs and satisfactions: A global survey. *Public opinion quarterly*, 40(4):459–467, 1976.

[148] Dana Rose Garfin, Roxane Cohen Silver, and E Alison Holman. The novel coronavirus (covid-2019) outbreak: Amplification of public health consequences by media exposure. *Health psychology*, 39(5):355, 2020.

[149] Lina Gega, Isaac Marks, and David Mataix-Cols. Computer-aided cbt self-help for anxiety and depressive disorders: Experience of a london clinic and future directions. *Journal of Clinical Psychology*, 60(2):147–57, 2004. doi: 10.1002/jclp.10241.

[150] AJ Gelenberg, MP Freeman, JC Markowitz, JF Rosenbaum, ME Thase, MH Trivedi, and RS Van Rhoads. American psychiatric association practice guidelines for the treatment of patients with major depressive disorder. *Am J Psychiatry*, 167(Suppl. 10):9–118, 2010. doi: 10.1176/appi.books.9780890423387.654001.

[151] Itzhak Gilat, Yishai Tobin, and Golan Shahar. Offering support to suicidal individuals in an online support group. *Archives of Suicide Research : Official Journal of the International Academy for Suicide Research*, 15(3):195–206, 2011. doi: 10.1080/13811118.2011.589675.

[152] Jeremy Ginsberg, Matthew H Mohebbi, Rajan S Patel, Lynnette Brammer, Mark S Smolinski, and Larry Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–1014, 2009.

[153] Kristina Gligorić, Arnaud Chiolero, Emre Kıcıman, Ryen W White, and Robert West. Population-scale dietary interests during the covid-19 pandemic. *Nature communications*, 13(1):1–14, 2022.

[154] Krystyna Glinski and Andrew C Page. Modifiability of neuroticism, extraversion, and agreeableness by group cognitive behaviour therapy for social anxiety disorder. *Behaviour Change*, 27(1):42–52, 2010.

[155] Melinda Goldner. How health status impacts the types of information consumers seek online. *Information, Community & Society*, 9(6):693–713, 2006.

[156] Friedrich M Götz, Stefan Stieger, Samuel D Gosling, Jeff Potter, and Peter J Rentfrow. Physical topography is associated with human personality. *Nature human behaviour*, 4(11):1135–1144, 2020.

[157] Christian Graff. Expressing relative differences (in percent) by the difference of natural logarithms. *Journal of Mathematical Psychology*, 60:82–85, 2014.

[158] William G Graziano and Nancy Eisenberg. Agreeableness: A dimension of personality. In *Handbook of personality psychology*, pages 795–824. Elsevier, 1997.

[159] Alessandro Grecucci and Alan G Sanfey. Emotion regulation and decision making. *Handbook of emotion regulation*, 2, 2014.

[160] Jose De Gregorio and Jong-Wha Lee. Education and income inequality: new evidence from cross-country data. *Review of income and wealth*, 48(3):395–416, 2002.

[161] James J Gross and Oliver P John. Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *Journal of personality and social psychology*, 85(2):348, 2003.

[162] Joseph G Grzywacz, David M Almeida, and Daniel A McDonald. Work–family spillover and daily reports of work and family stress in the adult labor force. *Family relations*, 51(1):28–36, 2002.

[163] Sumedha Gupta, Laura Montenovo, Thuy Dieu Nguyen, Felipe Lozano-Rojas, Ian M Schmutte, Kosali Ilayperuma Simon, Bruce A Weinberg, and Coady Wing. Effects of social distancing policy on labor market outcomes. *NBER Working paper*, (w27280), 2020.

[164] Timothy M Hale, Shelia R Cotten, Patricia Drentea, and Melinda Goldner. Rural-urban differences in general and health-related Internet use. *American Behavioral Scientist*, 53(9):1304–1325, 2010.

[165] Timothy M Hale, Melinda Goldner, Mike Stern, Patricia Drentea, and Shelia R Cotten. Patterns of online health searching 2002–2010: Implications for social capital, health disparities and the de-professionalization of medical knowledge. In *Technology, communication, disparities and government options in health and health care services*. Emerald Group Publishing Limited, 2014.

[166] Amanda K Hall, Jay M Bernhardt, Virginia Dodd, and Morgan W Vollrath. The digital health divide: evaluating online health information access and use among older adults. *Health Education & Behavior*, 42(2):202–209, 2015.

[167] Rebecca Hamilton. Scarcity and coronavirus. *Journal of Public Policy & Marketing*, 40(1):99–100, 2021.

[168] Sheila L Hammer, Karen Clark, Marcia Grant, and Matthew J Loscalzo. Seventeen years of progress for supportive care services: A resurvey of national cancer institute–designated comprehensive cancer centers. *Palliative & Supportive Care*, 13(4):917–925, 2015. doi: 10.1017/s1478951514000601.

[169] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. A systematic review of just-in-time adaptive interventions (jitais) to promote physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, 16(1):31, 2019. doi: 10.1186/s12966-019-0792-7.

[170] Wendy Hardeman, Julie Houghton, Kathleen Lane, Andy Jones, and Felix Naughton. A systematic review of just-in-time adaptive interventions (jitais) to promote physical activity. *International Journal of Behavioral Nutrition and Physical Activity*, 16(1):1–21, 2019.

[171] R Harding, E Epiphaniou, D Hamilton, S Bridger, V Robinson, R George, T Beynon, and IJ Higginson. What are the perceived needs and challenges of informal caregivers in home cancer palliative care? qualitative data to construct a feasible psycho-educational intervention. *Supportive Care in Cancer*, 20(9): 1975–1982, 2012. doi: 10.1007/s00520-011-1300-z.

[172] Eszter Hargittai. Second-level digital divide: Mapping differences in people's online skills. *arXiv preprint cs/0109068*, 2001.

[173] Eszter Hargittai. Digital na(t)ives? Variation in internet skills and uses among members of the "net generation". *Sociological inquiry*, 80(1):92–113, 2010.

[174] Eszter Hargittai and Amanda Hinnant. Digital inequality: Differences in young adults' use of the Internet. *Communication research*, 35(5):602–621, 2008.

[175] Eszter Hargittai and Aaron Shaw. Mind the skills gap: the role of Internet know-how and gender in differentiated contributions to Wikipedia. *Information, communication & society*, 18(4):424–442, 2015.

[176] Richard Harte, Liam Glynn, Alejandro Rodríguez-Molinero, Paul MA Baker, Thomas Scharf, Leo R Quinlan, Gearóid ÓLaighin, et al. A human-centered design methodology to enhance the usability, human factors, and user experience of connected health systems: a three-phase methodology. *JMIR human factors*, 4(1):e5443, 2017.

[177] Justin S Hatchimonji, Robert A Swendiman, Mark J Seamon, and Michael L

Nance. Trauma does not quarantine: violence during the covid-19 pandemic. *Annals of surgery*, 272(2):e53, 2020.

[178] Gillian R Hayes, Gregory D Abowd, John S Davis, Marion L Blount, Maria Ebling, and Elizabeth D Mynatt. Opportunities for pervasive computing in chronic cancer care. In *International Conference on Pervasive Computing*, pages 262–279. Springer, 2008. doi: 10.1007/978-3-540-79576-6_16.

[179] Sylvia Helbig and Lydia Fehm. Problems with homework in cbt: Rare exception or rather frequent? *Behavioural and cognitive psychotherapy*, 32(3):291–301, 2004.

[180] Ellen Johanna Helsper. A corresponding fields model for the links between social and digital exclusion. *Communication theory*, 22(4):403–426, 2012.

[181] Katherine E Henson, Rachael Brock, James Charnock, Bethany Wickramasinghe, Olivia Will, and Alexandra Pitman. Risk of suicide after cancer diagnosis in england. *JAMA Psychiatry*, 76(1):51–60, 2019. doi: 10.1001/jamapsychiatry.2018.3181.

[182] Miguel A Hernán and James M Robins. Using big data to emulate a target trial when a randomized trial is not available. *American journal of epidemiology*, 183 (8):758–764, 2016.

[183] Benjamin Herold. The disparities in remote learning under coronavirus (in charts). *Education Week*, 10, 2020.

[184] Susana García Herrero, Miguel Ángel Mariscal Saldaña, Javier García Rodriguez, and Dale O Ritzel. Influence of task demands on occupational stress: Gender differences. *Journal of Safety Research*, 43(5-6):365–374, 2012.

[185] Lisa M Hess and Kathleen C Insel. Chemotherapy-related change in cognitive function: A conceptual model. In *Oncology Nursing Forum*, volume 34, 2007. doi: 10.1188/07.onf.981-994.

[186] Nis Hjortskov, Dag Rissén, Anne Katrine Blangsted, Nils Fallentin, Ulf Lundberg, and Karen Søgaard. The effect of mental stress on heart rate variability and blood pressure during computer work. *European journal of applied physiology*, 92(1):84–89, 2004.

[187] Daniel E Ho, Kosuke Imai, Gary King, Elizabeth A Stuart, et al. MatchIt: nonparametric preprocessing for parametric causal inference. *J Stat Softw*, 42 (8):1–28, 2011.

[188] Frederick K Ho, Carlos A Celis-Morales, Stuart R Gray, S Vittal Katikireddi, Claire L Niedzwiedz, Claire Hastie, Lyn D Ferguson, Colin Berry, Daniel F Mackay, Jason MR Gill, et al. Modifiable and non-modifiable risk factors for covid-19, and comparison to risk factors for influenza and pneumonia: results from a uk biobank prospective cohort study. *BMJ open*, 10(11):e040402, 2020.

[189] Victoria Hollis, Artie Konrad, Aaron Springer, Matthew Antoun, Christopher Antoun, Rob Martin, and Steve Whittaker. What does all this data mean for my future mood? actionable analytics and targeted reflection for emotional well-being. *Human–Computer Interaction*, 32(5-6):208–267, 2017. doi: 10.1080/07370024.2016.1277724.

[190] Esther Howe, Jina Suh, Mehrab Bin Morshed, Daniel McDuff, Kael Rowan, Javier Hernandez, Marah Ihab Abdin, Gonzalo Ramos, Tracy Tran, and Mary P Czerwinski. Design of digital workplace stress-reduction intervention systems: Effects of intervention type and timing. In *CHI Conference on Human Factors in Computing Systems*, pages 1–16, 2022.

[191] Anna Huguet, Sanjay Rao, Patrick J Mcgrath, Lori Wozney, Mike Wheaton, Jill Conrod, and Sharlene Rozario. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLOS One*, 11(5):E0154248, 2016. doi: 10.1371/journal.pone.0154248.

[192] Anna Huguet, Sanjay Rao, Patrick J Mcgrath, Lori Wozney, Mike Wheaton, Jill Conrod, and Sharlene Rozario. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. *PLOS One*, 11(5):E0154248, 2016. URL https://doi.org/10.1371/journal.pone.0154248.

[193] Enid M Hunkeler, Wayne Katon, Lingqi Tang, John W Williams, Kurt Kroenke, Elizabeth HB Lin, Linda H Harpole, Patricia Arean, Stuart Levine, Lydia M Grypma, et al. Long term outcomes from the impact randomised trial for depressed elderly patients in primary care. *British Medical Journal*, 332(7536): 259–263, 2006. doi: 10.1136/bmj.38683.710255.be.

[194] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 17–24, 2003.

[195] Iana Roumenova Ianakieva. *A Task Analysis of Therapeutic Engagement in a Professionally Facilitated Online Intervention for Young Couples Affected by Breast Cancer*. PhD thesis, York University Toronto, ON, 2015. URL http://hdl.handle.net/10315/30677.

[196] Lucas D Introna. Workplace surveillance, privacy and distributive justice. *Acm Sigcas Computers and Society*, 30(4):33–39, 2000.

[197] Kelly E Irwin, Elyse R Park, Lauren E Fields, Amy E Corveleyn, Joseph A Greer,

Giselle K Perez, Catherine A Callaway, Jamie M Jacobs, Andrew A Nierenberg, Jennifer S Temel, et al. Bridge: Person-centered collaborative care for patients with serious mental illness and cancer. *The Oncologist*, 24(7):901–910, 2019. doi: 10.1634/theoncologist.2018-0488.

[198] Maia Jacobs, James Clawson, and Elizabeth D Mynatt. A cancer journey framework: Guiding the design of holistic health technology. In *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 114–121, 2016. doi: 10.4108/eai.16-5-2016.2263333.

[199] Maia L Jacobs, James Clawson, and Elizabeth D Mynatt. Articulating a patient-centered design space for cancer journeys. *EAI Endorsed Transactions on Pervasive Health and Technology*, 3(9), 2017. doi: 10.4108/eai.21-3-2017.152394.

[200] Paul B Jacobsen, Jimmie C Holland, and David P Steensma. Caring for the whole patient: The science of psychosocial care. American Society of Clinical Oncology, 2012. doi: 10.1200/jco.2011.41.4078.

[201] Luis G Jaimes, Martin Llofriu, and Andrew Raij. Preventer, a selection mechanism for just-in-time preventive interventions. *IEEE Transactions on Affective Computing*, 7(3):243–257, 2015. doi: 10.1109/TAFFC.2015.2490062.

[202] John Paul Jameson and Michael B Blank. Diagnosis and treatment of depression and anxiety in rural and nonrural primary care: National survey results. *Psychiatric Services*, 61(6):624–627, 2010. doi: 10.1176/ps.2010.61.6.624.

[203] Bernard J Jansen and Amanda Spink. How are we searching the world wide web? a comparison of nine search engine transaction logs. *Information processing & management*, 42(1):248–263, 2006.

[204] Abdulrahman Jbaily, Xiaodan Zhou, Jie Liu, Ting-Hwan Lee, Leila Kamareddine,

Stéphane Verguet, and Francesca Dominici. Air pollution exposure disparities across us population and income groups. *Nature*, 601(7892):228–233, 2022.

[205] Oliver P. John and Sanjay Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. 1999.

[206] Adrienne S Juarascio, Megan N Parker, Madeline A Lagacey, and Kathryn M Godfrey. Just-in-time adaptive interventions: A novel approach for enhancing skill utilization and acquisition in cognitive behavioral therapy for eating disorders. *International Journal of Eating Disorders*, 51(8):826–830, 2018.

[207] Ronald M Kadden and Mark D Litt. The role of self-efficacy in the treatment of substance use disorders. *Addictive behaviors*, 36(12):1120–1126, 2011.

[208] Jonathan W Kanter, Rachel C Manos, William M Bowe, David E Baruch, Andrew M Busch, and Laura C Rusch. What is behavioral activation?: A review of the empirical literature. *Clinical psychology review*, 30(6):608–620, 2010.

[209] Karim Sadik Kassam. *Assessment of emotional experience through facial expression*. Harvard University, 2010.

[210] Jay S Kaufman, Richard S Cooper, and Daniel L McGee. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology*, pages 621–628, 1997.

[211] John A Kaufman, Melvin D Livingston, and Kelli A Komro. Unemployment insurance program accessibility and suicide rates in the United States. *Preventive Medicine*, 141:106318, 2020.

[212] Gary King, Richard Nielsen, Carter Coberley, James E Pope, and Aaron Wells. Comparative effectiveness of matching methods for causal inference. *Unpublished manuscript, Institute for Quantitative Social Science, Harvard University, Cambridge, MA*, 2011.

[213] Gail Kinman and Fiona Jones. Lay representations of workplace stress: What do people really mean when they say they are stressed? *Work & stress*, 19(2): 101–120, 2005.

[214] Stephen Kisely, Elizabeth Crowe, and David Lawrence. Cancer-related mortality in people with mental illness. *JAMA Psychiatry*, 70(2):209–217, 2013. doi: 10.1001/jamapsychiatry.2013.278.

[215] David Kissane. Beyond the psychotherapy and survival debate: The challenge of social disparity, depression and treatment adherence in psychosocial cancer care. *Psycho-Oncology: Journal of the Psychological, Social and Behavioral Dimensions of Cancer*, 18(1):1–5, 2009. doi: 10.1002/pon.1493.

[216] Predrag Klasnja, Andrea Civan Hartzler, Kent T Unruh, and Wanda Pratt. Blowing in the wind: Unanchored patient information work during cancer care. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 193–202, 2010. doi: 10.1145/1753326.1753355.

[217] Predrag Klasnja, Shawna Smith, Nicholas J Seewald, Andy Lee, Kelly Hall, Brook Luers, Eric B Hekler, and Susan A Murphy. Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of heartsteps. *Annals of Behavioral Medicine*, 53(6):573–582, 2019.

[218] Artie Konrad, Victoria Bellotti, Nicole Crenshaw, Simon Tucker, Les Nelson, Honglu Du, Peter Pirolli, and Steve Whittaker. Finding the adaptive sweet spot: Balancing compliance and achievement in automated stress reduction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 3829–3838, 2015.

[219] Emily Kontos, Kelly D Blake, Wen-Ying Sylvia Chou, and Abby Prestin. Predictors of eHealth usage: insights on the digital divide from the Health Information

National Trends Survey 2012. *Journal of Medical Internet Research*, 16(7):e3117, 2014.

[220] Pål Kraft, Filip Drozd, and Elin Olsen. Digital therapy: Addressing willpower as part of the cognitive-affective processing system in the service of habit change. In *International Conference on Persuasive Technology*, pages 177–188. Springer, 2008. doi: 10.1007/978-3-540-68504-3_16.

[221] AMH Krebber, LM Buffart, G Kleijn, IC Riepma, R De Bree, CR Leemans, A Becker, J Brug, A Van Straten, P Cuijpers, et al. Prevalence of depression in cancer patients: a meta-analysis of diagnostic interviews and self-report instruments. *Psycho-Oncology*, 23(2):121–130, 2014. doi: 10.1002/pon.3409.

[222] Kurt Kroenke and Jurgen Unutzer. Closing the False Divide: Sustainable Approaches to Integrating Mental Health Services into Primary Care. *Journal of General Internal Medicine*, 32(4):404–410, 2017. ISSN 15251497. doi: 10.1007/s11606-016-3967-9.

[223] Kurt Kroenke, Robert L Spitzer, and Janet BW Williams. The phq-9: Validity of a brief depression severity measure. *Journal of General Internal medicine*, 16(9): 606–613, 2001. doi: 10.1046/j.1525-1497.2001.016009606.x.

[224] Kurt Kroenke, Dale Theobald, Jingwei Wu, Kelli Norton, Gwendolyn Morrison, Janet Carpenter, and Wanzhu Tu. Effect of telecare management on pain and depression in patients with cancer: A randomized trial. *Journal of the American Medical Association*, 304(2):163–171, 2010. doi: 10.1001/jama.2010.944.

[225] Eric Kuhn, Nitya Kanuri, Julia E Hoffman, Donn W Garvert, Josef I Ruzek, and C Barr Taylor. A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. *Journal of consulting and clinical psychology*, 85(3):267, 2017.

[226] Peter Kuhn and Hani Mansour. Is internet job search still ineffective? *The Economic Journal*, 124(581):1213–1233, 2014.

[227] Kostadin Kushlev, Jason Proulx, and Elizabeth W Dunn. " silence your phones" smartphone notifications increase inattention and hyperactivity symptoms. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 1011–1020, 2016.

[228] Vasileios Lampos, Maimuna S Majumder, Elad Yom-Tov, Michael Edelstein, Simon Moura, Yohhei Hamada, Molebogeng X Rangaka, Rachel A McKendry, and Ingemar J Cox. Tracking COVID-19 using online search. *NPJ digital medicine*, 4(1):1–11, 2021.

[229] Michelle S Landwehr, Samantha E Watson, Catherine F Macpherson, Katherine A Novak, and Rebecca H Johnson. The cost of cancer: A retrospective analysis of the financial impact of cancer on young adults. *Cancer Medicine*, 5 (5):863–870, 2016. doi: 10.1002/cam4.657.

[230] Øyvind Langsrud. Anova for unbalanced data: Use type ii instead of type iii sums of squares. *Statistics and Computing*, 13(2):163–167, 2003.

[231] Danielle C Lavallee, Kate E Chenok, Rebecca M Love, Carolyn Petersen, Erin Holve, Courtney D Segal, and Patricia D Franklin. Incorporating patient-reported outcomes into health care to engage patients and enhance care. *Health Affairs*, 35(4):575–582, 2016.

[232] Danielle C Lavallee, Jenney R Lee, Elizabeth Austin, Richard Bloch, Sarah O Lawrence, Debbe McCall, Sean A Munson, Mara B Nery-Hurwit, and Dagmar Amtmann. mhealth and patient generated health data: stakeholder perspectives on opportunities and barriers for transforming healthcare. *Mhealth*, 6, 2020.

[233] Jay L Lebow. The challenges of covid-19 for divorcing and post-divorce families. *Family process*, 59(3):967–973, 2020.

[234] Michael Lechner. The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends in Econometrics*, 4(3):165–224, 2011. ISSN 1551-3076. doi: 10.1561/0800000014. URL http://dx.doi.org/10.1561/0800000014.

[235] Reeva Lederman, Greg Wadley, John Gleeson, Sarah Bendall, and Mario Álvarez-Jiménez. Moderated online social therapy: Designing and evaluating technology for mental health. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 21(1):5, 2014. doi: 10.1145/2513179.

[236] Paul M Leonardi and Jeffrey W Treem. Behavioral visibility: A new paradigm for organization studies in the age of digitization, digitalization, and datafication. *Organization Studies*, 41(12):1601–1625, 2020.

[237] James R Lewis. The system usability scale: past, present, and future. *International Journal of Human–Computer Interaction*, 34(7):577–590, 2018.

[238] Madeline Li, Erin B Kennedy, Nelson Byrne, Caroline Gérin-Lajoie, Mark R Katz, Homa Keshavarz, Scott Sellick, and Esther Green. Systematic review and meta-analysis of collaborative care interventions for depression in patients with cancer. *Psycho-Oncology*, 26(5):573–587, 2017. doi: 10.1002/pon.4286.

[239] Weihua Li and Beth Schwartzapfel. Is domestic violence rising during the coronavirus shutdown? here's what the data shows. https://www.themarshallproject.org/2020/04/22/is-domestic-violence-rising-during-the-coronavirus-shutdown-here-s-what-the-data-shows/ (Accessed 2021-08-17), Apr 2020.

[240] Peng Liao, Predrag Klasnja, Ambuj Tewari, and Susan A Murphy. Sample size

calculations for micro-randomized trials in mhealth. *Statistics in medicine*, 35 (12):1944–1971, 2016.

[241] Chen-Tan Lin, Loretta Wittevrongel, Laurie Moore, Brenda L Beaty, and Stephen E Ross. An internet-based patient-provider communication system: Randomized controlled trial. *Journal of Medical Internet Research*, 7(4):e47, 2005. doi: 10.2196/jmir.7.4.e47.

[242] Yu-Hsuan Lin, Chun-Hao Liu, and Yu-Chuan Chiu. Google searches for the keywords of "wash hands" predict the speed of national spread of covid-19 outbreak among 21 countries. *Brain, behavior, and immunity*, 87:30–32, 2020.

[243] Jake Linardon and Matthew Fuller-Tyszkiewicz. Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of consulting and clinical psychology*, 88(1):1, 2020.

[244] Oliver Lindhiem, Charles B Bennett, Dana Rosen, and Jennifer Silk. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. *Behavior modification*, 39(6):785–804, 2015.

[245] Marsha Linehan. *DBT Skills training manual.* Guilford Publications, 2014.

[246] Xin Liu, Josh Fromm, Shwetak Patel, and Daniel McDuff. Multi-task temporal shift attention networks for on-device contactless vitals measurement. *NeurIPS*, 2020.

[247] Zijian Long, Rajwa Alharthi, and Abdulmotaleb El Saddik. Needfull–a tweet analysis platform to study human needs during the covid-19 pandemic in new york state. *Ieee Access*, 8:136046–136055, 2020.

[248] Peter F Lovibond and Sydney H Lovibond. The structure of negative emotional states: Comparison of the depression anxiety stress scales (dass) with the beck

depression and anxiety inventories. *Behaviour research and therapy*, 33(3): 335–343, 1995.

[249] Bernd Löwe, Kurt Kroenke, and Kerstin Gräfe. Detecting and monitoring depression with a two-item questionnaire (phq-2). *Journal of Psychosomatic Research*, 58(2):163–171, 2005. doi: 10.1016/j.jpsychores.2004.09.006.

[250] Gale M Lucas, Jonathan Gratch, Aisha King, and Louis-Philippe Morency. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37:94–100, 2014. doi: 10.1016/j.chb.2014.04.043.

[251] Anita Lungu and Marsha M Linehan. Dialectical behavior therapy: A comprehensive multi-and transdiagnostic intervention. *The Oxford handbook of cognitive and behavioral therapies*, pages 200–214, 2016.

[252] Aaron R. Lyon and Kelly Koerner. User-Centered Design for Psychosocial Intervention Development and Implementation. *Clinical Psychology: Science and Practice*, 23(2):180–200, 2016. ISSN 14682850. doi: 10.1111/cpsp.12154.

[253] Aaron R Lyon, Sean A Munson, Brenna N Renn, David C Atkins, Michael D Pullmann, Emily Friedman, and Patricia A Areán. Use of Human-Centered Design to Improve Implementation of Evidence-Based Psychotherapies in Low-Resource Communities: Protocol for Studies Applying a Framework to Assess Usability . *JMIR Research Protocols*, 8(10):e14990, 2019. ISSN 1929-0748. doi: 10.2196/14990.

[254] GC Turnbull Macdonald, F Baldassarre, P Brown, J Hatton-Bauer, M Li, E Green, and S Lebel. Psychosocial care for cancer: A framework to guide practice, and actionable recommendations for ontario. *Current Oncology*, 19(4):209, 2012. doi: 10.3747/co.19.981.

[255] Elisabeth Mahase. Covid-19: EU states report 60% rise in emergency calls about domestic violence. *BMJ: British Medical Journal (Online)*, 369, 2020.

[256] Tina Malti, Gil G. Noam, Andreas Beelmann, and Simon Sommer. Toward dynamic adaptation of psychological interventions for child and adolescent development and mental health. *Journal of Clinical Child and Adolescent Psychology*, 45(6):827–836, 2016. ISSN 15374416. doi: 10.1080/15374416.2016.1239539. URL http://dx.doi.org/10.1080/15374416.2016.1239539.

[257] Lydia Manikonda and Munmun De Choudhury. Modeling and understanding visual attributes of mental health disclosures in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2017. doi: 10.1145/3025453.3025932.

[258] Sonja March, Jamin Day, Gabrielle Ritchie, Arlen Rowe, Jeffrey Gough, Tanya Hall, Chin Yan Jackie Yuen, Caroline Leanne Donovan, Michael Ireland, et al. Attitudes toward e-mental health services in a community sample of adults: online survey. *Journal of medical Internet research*, 20(2):e9109, 2018.

[259] Sue M Marcus, Juned Siddique, Thomas R Ten Have, Robert D Gibbons, Elizabeth Stuart, and Sharon-Lise T Normand. Balancing treatment comparisons in longitudinal studies. *Psychiatric Annals*, 38(12), 2008.

[260] Gloria Mark, Yiran Wang, and Melissa Niiya. Stress and multitasking in everyday college life: an empirical study of online activity. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 41–50, 2014.

[261] Gloria Mark, Shamsi T Iqbal, Mary Czerwinski, Paul Johns, Akane Sano, and Yuliya Lutchyn. Email duration, batching and self-interruption: Patterns of email use on productivity and stress. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, pages 1717–1728. ACM, 2016.

[262] Gloria Mark, Shamsi Iqbal, and Mary Czerwinski. How blocking distractions affects workplace focus and productivity. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pages 928–934, 2017.

[263] Ron Martin and Peter Sunley. On the notion of regional economic resilience: conceptualization and explanation. *Journal of economic geography*, 15(1):1–42, 2015.

[264] Abraham H. Maslow. A theory of human motivation. *Psychol. Rev.*, 1943.

[265] Abraham H. Maslow. New introduction: Religions, values, and peak-experiences. *Journal of Transpersonal Psychology*, 2(2), 1970.

[266] Eugene W Mathes and Linda L Edwards. An empirical test of maslow's theory of motivation. *Journal of Humanistic Psychology*, 18(1):75–77, 1978.

[267] Mark Matthews, Gavin Doherty, David Coyle, and John Sharry. Designing mobile applications to support mental health interventions. In *Handbook of research on user interface design and evaluation for mobile technology*, pages 635–656. IGI Global, 2008. doi: 10.4018/978-1-59904-871-0.ch038.

[268] Brent T Mausbach, Philip Yeung, Taylor Bos, and Scott A Irwin. Health care costs of depression in patients diagnosed with cancer. *Psycho-Oncology*, 27(7): 1735–1741, 2018. doi: 10.1002/pon.4716.

[269] Manfred A Max-Neef. Human scale development: conception, application and further reflections. 1991.

[270] Leah M Mayo and Markus Heilig. In the face of stress: Interpreting individual differences in stress-induced facial expressions. *Neurobiology of stress*, 10:100166, 2019.

[271] Robert R McCrae and Paul T Costa Jr. Conceptions and correlates of openness to experience. In *Handbook of personality psychology*, pages 825–847. Elsevier, 1997.

[272] Robert R McCrae and Oliver P John. An introduction to the five-factor model and its applications. *Journal of personality*, 60(2):175–215, 1992.

[273] Margaret V McDonald, Steven D Passik, William Dugan, Barry Rosenfeld, Dale E Theobald, and Sara Edgerton. Nurses' recognition of depression in their patients with cancer. In *Oncology Nursing Forum*, volume 26, pages 593–599, 1999. doi: 10.1002/pon.3191.

[274] Daniel McDuff, Eunice Jun, Kael Rowan, and Mary Czerwinski. Longitudinal observational evidence of the impact of emotion regulation strategies on affective expression. *IEEE Transactions on Affective Computing*, 2019.

[275] Daniel McDuff, Kael Rowan, Piali Choudhury, Jessica Wolk, ThuVan Pham, and Mary Czerwinski. A multimodal emotion sensing platform for building emotion-aware applications. *arXiv preprint arXiv:1903.12133*, 2019.

[276] Rick McHugh and Will Kimball. How Low Can We Go? State Unemployment Insurance Programs Exclude Record Numbers of Jobless Workers. *EPI, March*, 9, 2015.

[277] Anna Meijer, Michelle Roseman, Katherine Milette, James C Coyne, Michael E Stefanek, Roy C Ziegelstein, Erin Arthurs, Allison Leavens, Steven C Palmer, Donna E Stewart, et al. Depression screening and patient outcomes in cancer: A systematic review. *PLoS One*, 6(11):e27181, 2011. doi: 10.1371/journal.pone.0027181.

[278] Marijke Melles, Armagan Albayrak, and Richard Goossens. Innovating health

care: key characteristics of human-centered design. *International Journal for Quality in Health Care*, 33(Supplement_1):37–44, 2021.

[279] Michael Mikolasek, Claudia M Witt, Jürgen Barth, et al. Adherence to a mindfulness and relaxation self-care app for cancer patients: mixed-methods feasibility study. *JMIR mHealth and uHealth*, 6(12):e11271, 2018.

[280] Stephanie Misono, Noel S Weiss, Jesse R Fann, Mary Redman, and Bevan Yueh. Incidence of suicide in persons with cancer. *Journal of Clinical Oncology*, 26 (29):4731, 2008. doi: 10.1200/jco.2007.13.8941.

[281] Vance F Mitchell and Pravin Moudgill. Measurement of maslow's need hierarchy. *Organizational Behavior and Human Performance*, 16(2):334–349, 1976.

[282] David C Mohr, Michelle Nicole Burns, Stephen M Schueller, Gregory Clarke, and Michael Klinkman. Behavioral intervention technologies: Evidence review and recommendations for future research in mental health. *General Hospital Psychiatry*, 35(4):332–338, 2013. doi: 10.1016/j.genhosppsych.2013.03.008.

[283] David C Mohr, Enid Montague, Colleen Stiles-Shields, Susan M Kaiser, Christopher Brenner, Eric Carty-Fickes, Hannah Palac, and Jenna Duffecy. Medlink: A mobile intervention to address failure points in the treatment of depression in general medicine. In *Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare*, pages 100–107. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, 2015. doi: 10.4108/icst.pervasivehealth.2015.259042.

[284] David C Mohr, Kathryn Noth Tomasino, Emily G Lattie, Hannah L Palac, Mary J Kwasny, Kenneth Weingardt, Chris J Karr, Susan M Kaiser, Rebecca C Rossom, Leland R Bardsley, Lauren Caccamo, Colleen Stiles-Shields, and Stephen M Schueller. Intellicare: An eclectic, skills-based app suite for the treatment of

depression and anxiety. *Journal of Medical Internet Research*, 19(1), 2017. URL https://dx.doi.org/10.2196%2Fjmir.6645.

[285] David C Mohr, Francisca Azocar, Andrew Bertagnolli, Tanzeem Choudhury, Paul Chrisp, Richard Frank, Henry Harbin, Trina Histon, Debra Kaysen, Camille Nebeker, et al. Banbury forum consensus statement on the path forward for digital mental health treatment. *Psychiatric Services*, 72(6):677–683, 2021.

[286] Nathalie Moise, Ravi N Shah, Susan Essock, Amy Jones, Jay Carruthers, Margaret A Handley, Lauren Peccoralo, and Lloyd Sederer. Sustainability of collaborative care management for depression in primary care settings with academic affiliations across new york state. *Implementation Science*, 13(1):128, 2018. doi: 10.1186/s13012-018-0818-6.

[287] Ramin Mojtabai, Mark Olfson, Nancy a Sampson, Benjamin Druss, Philip S Wang, Kenneth B Wells, Harold a Pincus, and Ronald C Kessler. Barriers to mental health treatment: results from the WHO World Mental Health surveys. *Psychological Medicine*, 41(8):1751–1761, 2011. ISSN 0033-2917. doi: 10.1017/S0033291710002291.Barriers.

[288] Catherine E Mosher, Katherine N DuHamel, Christine Rini, Geoffrey Corner, Joanne Lam, and William H Redd. Quality of life concerns and depression among hematopoietic stem cell transplant survivors. *Supportive Care in Cancer*, 19(9):1357–1365, 2011. doi: 10.1007/s00520-010-0958-y.

[289] Stuart Moss et al. An introduction to the entertainment industry. *The entertainment industry: An introduction*, pages 1–18, 2010.

[290] Sarah Myruski, Samantha Denefrio, and Tracy A Dennis-Tiwary. Stress and emotion regulation. *The oxford handbook of stress and mental Health*, page 415, 2019.

[291] Inbal Nahum-Shani, Shawna N Smith, Ambuj Tewari, Katie Witkiewitz, Linda M Collins, Bonnie Spring, and S Murphy. Just in time adaptive interventions (jitais): An organizing framework for ongoing health behavior support. *Methodology Center technical report*, 2014:14–126, 2014. doi: 10.1007/s12160-016-9830-8.

[292] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. Just-in-time adaptive interventions (jitais) in mobile health: key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine*, 52(6):446–462, 2018.

[293] Ron Neal and Joanne Sibug. Workplace Stress on the Rise With 83% of Americans Frazzled by Something at Work. https://www.globenewswire.com/news-release/2013/04/09/536945/10027728/en/Workplace-Stress-on-the-Rise-With-83-of-Americans-Frazzled-by-Something-at-Work.html (Accessed 2021-08-20).

[294] Martha Neary and Stephen M Schueller. State of the field of mental health apps. *Cognitive and Behavioral Practice*, 25(4):531–537, 2018.

[295] National Comprehensive Cancer Network et al. Distress management. clinical practice guidelines. *Journal of the National Comprehensive Cancer Network: JNCCN*, 1(3):344, 2003. doi: 10.6004/jnccn.2003.0031.

[296] Ryan D Nipp, Areej El-Jawahri, Samantha M Moran, Sara M D'arpino, P Connor Johnson, Daniel E Lage, Risa L Wong, William F Pirl, Lara Traeger, Inga T Lennes, et al. The relationship between physical and psychological symptoms and health care utilization in hospitalized patients with advanced cancer. *Cancer*, 123(23):4720–4727, 2017. doi: 10.1002/cncr.30912.

[297] Laurel Northouse, Anna-leila Williams, Barbara Given, and Ruth McCorkle.

Psychosocial care for family caregivers of patients with cancer. *Journal of Clinical Oncology*, 30(11):1227–1234, 2012. doi: 10.1200/jco.2011.39.5798.

[298] National Institutes of Health: National Institute of Mental Health. Any mental illness (ami) among u.s. adults, 2015. URL https://www.nimh.nih.gov/health/statistics/prevalence/any-mental-illness-ami-among-us-adults.shtml.

[299] Mark Olfson, Ramin Mojtabai, Nancy A Sampson, Irving Hwang, Benjamin Druss, Philip S Wang, Kenneth B Wells, Harold Alan Pincus, and Ronald C Kessler. Dropout from outpatient mental health care in the united states. *Psychiatric Services*, 60(7):898–907, 2009. URL https://doi.org/10.1176/ps.2009.60.7.898.

[300] Nosayaba Osazuwa-Peters, Matthew C. Simpson, Longwen Zhao, Eric Adjei Boakye, Stephanie I. Olomukoro, Teresa Deshields, Travis M. Loux, Mark A. Varvares, and Mario Schootman. Suicide risk among cancer survivors: Head and neck versus other cancers. *Cancer*, 124(20):4072–4079, 2018. doi: 10.1002/cncr.31675. URL https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.1002/cncr.31675.

[301] Thomas P. O'Toole, P. A. Pirraglia, D. Dosa, C. Bourgault, S. Redihan, M. B. O'Toole, J. Blumen, and Primary Care-Special Populations Treatment Team. Building care systems to improve access for high-risk and vulnerable veteran populations. *Journal of general internal medicine*, 26 Suppl 2:683–688, 2011. ISSN 15251497. doi: 10.1007/s11606-011-1818-2.

[302] Ann EK Page, Nancy E Adler, et al. *Cancer Care for the Whole Patient: Meeting Psychosocial Health Needs.* National Academies Press, 2008. doi: 10.17226/11993.

[303] Sandro Pampallona, P Bollini, G Tibaldi, B Kupelnick, and Carmine Munizza. Patient adherence in the treatment of depression. *The British Journal of Psychiatry*, 180(2):104–109, 2002. doi: 10.1192/bjp.180.2.104.

[304] John Paparrizos, Ryen W White, and Eric Horvitz. Screening for pancreatic adenocarcinoma using signals from web search logs: Feasibility study and results. *Journal of oncology practice*, 12(8):737–744, 2016.

[305] Pablo Paredes, Ran Gilad-Bachrach, Mary Czerwinski, Asta Roseway, Kael Rowan, and Javier Hernandez. Poptherapy: Coping with stress through pop-culture. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*, pages 109–117, 2014.

[306] Zachary Parolin and Emma K Lee. Large socio-economic, geographic and demographic disparities exist in exposure to school closures. *Nature human behaviour*, 5(4):522–528, 2021.

[307] Steven D Passik, William Dugan, Margaret V McDonald, Barry Rosenfeld, Dale E Theobald, and Sara Edgerton. Oncologists' recognition of depression in their patients with cancer. *Journal of Clinical Oncology*, 16(4):1594–1600, 1998. doi: 10.1200/JCO.1998.16.4.1594.

[308] Nilay Patel and Sophie Erickson. Remote learning is here to stay - can we make it better? https://www.theverge.com/21570482/remote-learning-khan-academy-interview-decoder-podcast (Accessed 2021-05-30), Nov 2020.

[309] Michael J Paul, Ryen W White, and Eric Horvitz. Search and breast cancer: On episodic shifts of attention over life histories of an illness. *ACM Transactions on the Web (TWEB)*, 10(2):1–27, 2016.

[310] Sachin R Pendse, Daniel Nkemelu, Nicola J Bidwell, Sushrut Jadhav, Soumitra Pathare, Munmun De Choudhury, and Neha Kumar. From treatment to healing: Envisioning a decolonial digital mental health. In *CHI Conference on Human Factors in Computing Systems*, pages 1–23, 2022.

[311] AT Pereira, C Cabacos, S Soares, A Araújo, A Manão, AP Amaral, R Sousa, and A Macedo. Relationship between personality, adherence to (mental) health behaviours and psychological distress during the covid-19 pandemic. *European Psychiatry*, 65(S1):S499–S500, 2022.

[312] Verónica Pérez-Rosas, Xuetong Sun, Christy Li, Yuchen Wang, Kenneth Resnicow, and Rada Mihalcea. Analyzing the quality of counseling conversations: the tell-tale signs of high-quality counseling. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, 2018.

[313] C Petersen and P DeMuro. Legal and regulatory considerations associated with use of patient-generated health data from social media and mobile health (mhealth) devices. *Applied clinical informatics*, 6(01):16–26, 2015.

[314] Betty Pfefferbaum and Carol S North. Mental health and the covid-19 pandemic. *New England Journal of Medicine*, 383(6):510–512, 2020.

[315] Errol J Philip, Thomas V Merluzzi, Zhiyong Zhang, and Carolyn A Heitzmann. Depression and cancer survivorship: Importance of coping self-efficacy in post-treatment survivors. *Psycho-Oncology*, 22(5):987–994, 2013. doi: 10.1002/pon.3088.

[316] William F Pirl, Jesse R Fann, Joseph A Greer, Ilana Braun, Teresa Deshields, Caryl Fulcher, Elizabeth Harvey, Jimmie Holland, Vicki Kennedy, Mark Lazenby, et al. Recommendations for the implementation of distress screening programs in cancer centers: Report from the american psychosocial oncology society (apos), association of oncology social work (aosw), and oncology nursing society (ons) joint task force. *Cancer*, 120(19):2946–2954, 2014. doi: 10.1002/cncr.28750.

[317] Alexandra Pitman, Sahil Suleman, Nicholas Hyde, and Andrew Hodgkiss. De-

pression and anxiety in patients with cancer. *Bmj*, 361:k1415, 2018. doi: 10.1136/bmj.k1415.

[318] Thane S Pittman and Kate R Zeigler. Basic human needs. 2007.

[319] Martin Prince, Vikram Patel, Shekhar Saxena, Mario Maj, Joanna Maselko, Michael R. Phillips, and Atif Rahman. No health without mental health. *Lancet*, 370(9590):859–877, 2007. ISSN 01406736. doi: 10.1016/S0140-6736(07)61238-0.

[320] James O Prochaska and Wayne F Velicer. The transtheoretical model of health behavior change. *American journal of health promotion*, 12(1):38–48, 1997.

[321] Hsiao-Tieh Pu, Shui-Lung Chuang, and Chyan Yang. Subject categorization of query terms for exploring web users' search interests. *Journal of the American Society for Information Science and Technology*, 53(8):617–630, 2002.

[322] Massimo Ragnedda and Maria Laura Ruiu. *Digital capital: A Bourdieusian perspective on the digital divide.* Emerald Group Publishing, 2020.

[323] J. Raifman, K. Nocka, D. Jones, J. Bor, S. Lipson, J. Jay, M. Cole, N. Krawczyk, P. Chan, S. Galea, et al. COVID-19 US state policy database. *Inter-university Consortium for Political and Social Research (ICPSR)*, 2020.

[324] Anita Ramsetty and Cristin Adams. Impact of the digital divide in the age of COVID-19. *Journal of the American Medical Informatics Association*, 27(7): 1147–1148, 2020.

[325] Gurvaneet S Randhawa, David K Ahern, and Bradford W Hesse. Information technology-enabled team-based, patient-centered care: The example of depression screening and management in cancer care. *Health Policy and Technology*, 6 (1):67–71, 2017. doi: 10.1016/j.hlpt.2016.08.001.

[326] Lori Raney, David Bergman, John Torous, and Michael Hasselberg. Digitally driven integrated primary care and behavioral health: How technology can expand access to effective treatment. *Current Psychiatry Reports*, 19(11):86, 2017. doi: 10.1007/s11920-017-0838-y.

[327] Michael Ratcliffe, Charlynn Burd, Kelly Holder, and Alison Fields. Defining rural at the US Census Bureau. *American community survey and geography brief*, 1(8), 2016.

[328] Julie Ray. Americans' Stress, Worry and Anger Intensified in 2018. https://news.gallup.com/poll/249098/americans-stress-worry-anger-intensified-2018.aspx (Accessed 2021-08-20), 2019.

[329] Frederick F Reichheld. The one number you need to grow. *Harvard business review*, 81(12):46–55, 2003.

[330] James D Reschovsky and Andrea B Staiti. Access and quality: Does rural america lag behind? *Health Affairs*, 24(4):1128–1139, 2005. doi: 10.1377/hlthaff.24.4.1128.

[331] Julia Reynolds, Kathleen Griffiths, Helen Christensen, et al. Anxiety and depression: Online resources and management tools. *Australian Family Physician*, 40(6):382, 2011.

[332] Shirley Reynolds. Psychological well-being at work: is prevention better than cure? *Journal of psychosomatic research*, 43(1):93–102, 1997.

[333] Katherine M Richardson and Hannah R Rothstein. Effects of occupational stress management intervention programs: a meta-analysis. *Journal of occupational health psychology*, 13(1):69, 2008.

[334] Katherine Rieke, Kendra K. Schmid, William Lydiatt, Julia Houfek, Eugene Boilesen, and Shinobu Watanabe-Galloway. Depression and survival in head

and neck cancer patients. *Oral Oncology*, 65:76 – 82, 2017. ISSN 1368-8375. doi: https://doi.org/10.1016/j.oraloncology.2016.12.014.

[335] Shireen L Rizvi, Linda A Dimeff, Julie Skutch, David Carroll, and Marsha M Linehan. A pilot study of the dbt coach: An interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. *Behavior Therapy*, 42(4):589–600, 2011. doi: 10.1016/j.beth.2011.01.003.

[336] Shireen L Rizvi, Christopher D Hughes, and Marget C Thomas. The dbt coach mobile application as an adjunct to treatment for suicidal and self-injuring individuals with borderline personality disorder: A preliminary evaluation and challenges to client utilization. *Psychological Services*, 13(4):380–388, 2016. doi: 10.1037/ser0000100.

[337] Brent W Roberts, Jing Luo, Daniel A Briley, Philip I Chow, Rong Su, and Patrick L Hill. A systematic review of personality trait change through intervention. *Psychological Bulletin*, 143(2):117, 2017.

[338] Laura Robinson, Shelia R Cotten, Hiroshi Ono, Anabel Quan-Haase, Gustavo Mesch, Wenhong Chen, Jeremy Schulz, Timothy M Hale, and Michael J Stern. Digital inequalities and why they matter. *Information, communication & society*, 18(5):569–582, 2015.

[339] Everett M Rogers, Arvind Singhal, and Margaret M Quinlan. Diffusion of innovations. In *An integrated approach to communication theory and research*, pages 432–448. Routledge, 2014.

[340] Lloyd H Rogler. Methodological sources of cultural insensitivity in mental health research. *American psychologist*, 54(6):424, 1999.

[341] Bruce L Rollman, Bea Herbeck Belnap, Sati Mazumdar, Kaleab Z Abebe, Jordan F Karp, Eric J Lenze, and Herbert C Schulberg. Telephone-delivered stepped

collaborative care for treating anxiety in primary care: A randomized controlled trial. *Journal of General Internal Medicine*, 32(3):245–255, 2017. doi: 10.1007/s11606-016-3873-1.

[342] Paul R Rosenbaum and Donald B Rubin. The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55, 1983.

[343] Catherine E Ross and Chia-ling Wu. The links between education and health. *American sociological review*, pages 719–745, 1995.

[344] Stuart Ross and Sophie Aitken. "If It Hadn't Been Online I Don't Think I Would Have Applied": Applicant Experiences of an Online Family Violence Intervention Order Process. *Journal of interpersonal violence*, 2020.

[345] Andrew J Roth, Alice B Kornblith, Laure Batel-Copel, Elizabeth Peabody, Howard I Scher, and Jimmie C Holland. Rapid screening for psychologic distress in men with prostate carcinoma: a pilot study. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, 82(10):1904–1908, 1998. doi: 10.1002/(SICI)1097-0142(19980515)82:10<1904::AID-CNCR13>3.0.CO;2-X.

[346] Donald B Rubin. Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3):169–188, 2001.

[347] Benjamin J Ryan, Damon Coppola, Deon V Canyon, Mark Brickhouse, and Raymond Swienton. Covid-19 community stabilization and sustainability framework: an integration of the maslow hierarchy of needs and social determinants of health. *Disaster medicine and public health preparedness*, 14(5):623–629, 2020.

[348] Anas M Saad, Mohamed M Gad, Muneer J Al-Husseini, Mohamad A AlKhayat, Ahmad Rachid, Ahmad Samir Alfaar, and Hesham M Hamoda. Suicidal death

within a year of a cancer diagnosis: A population-based study. *Cancer*, 125(6): 972–979, 2019. doi: 10.1002/cncr.31876.

[349] Adam Sadilek, Stephanie Caty, Lauren DiPrete, Raed Mansour, Tom Schenk, Mark Bergtholdt, Ashish Jha, Prem Ramaswami, and Evgeniy Gabrilovich. Machine-learned epidemiology: real-time detection of foodborne illness at scale. *NPJ digital medicine*, 1(1):1–7, 2018.

[350] Koustuv Saha, Ted Grover, Stephen M. Mattingly, Vedant Das Swain, Pranshu Gupta, Gonzalo J. Martinez, Pablo Robles-Granda, Gloria Mark, Aaron Striegel, and Munmun De Choudhury. Person-Centered Predictions of Psychological Constructs with Social Media Contextualized by Multimodal Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5 (1):1–32, 2021. ISSN 24749567. doi: 10.1145/3448117.

[351] Michelle J Saksena, Abigail M Okrent, Tobenna D Anekwe, Clare Cho, Christopher Dicken, Anne Effland, Howard Elitzak, Joanne Guthrie, Karen S Hamrick, Jeffrey Hyman, et al. America's eating habits: food away from home. Technical report, United States Department of Agriculture, Economic Research Service, 2018.

[352] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, et al. Hci and affective health: Taking stock of a decade of studies and charting future research directions. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, page 245. ACM, 2019. doi: 10.1145/3290605.3300475.

[353] Mark Sanderson and Susan Dumais. Examining repetition in user search behavior. In *European Conference on Information Retrieval*, pages 597–604. Springer, 2007.

[354] Akane Sano, Paul Johns, and Mary Czerwinski. Healthaware: An advice system for stress, sleep, diet and exercise. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 546–552. IEEE, 2015. URL https://doi.org/10.1109/ACII.2015.7344623.

[355] Akane Sano, Paul Johns, and Mary Czerwinski. Designing opportune stress intervention delivery timing using multi-modal data. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 346–353. IEEE, 2017.

[356] Urmimala Sarkar, Andrew J Karter, Jennifer Y Liu, Nancy E Adler, Robert Nguyen, Andrea López, and Dean Schillinger. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. *Journal of the American Medical Informatics Association*, 18(3):318–321, 2011.

[357] Marc J. Schabracq, Jacques A.M. Winnubst, and Cary L. Cooper. The Handbook of Work and Health Psychology: Second Edition. *The Handbook of Work and Health Psychology: Second Edition*, pages 1–619, 2004. doi: 10.1002/0470013400.

[358] Jessica L Schleider and John R Weisz. Reducing risk for anxiety and depression in adolescents: Effects of a single-session intervention teaching that personality can change. *Behaviour research and therapy*, 87:170–181, 2016.

[359] Jessica Lee Schleider, Mallory Dobias, Jenna Sung, Emma Mumper, and Michael C Mullarkey. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health. *JMIR mental health*, 7(6):e20513, 2020.

[360] Eric C Schneider, Alan M Zaslavsky, and Arnold M Epstein. Use of high-cost

operative procedures by medicare beneficiaries enrolled in for-profit and not-for-profit health plans. *New England Journal of Medicine*, 350(2):143–150, 2004.

[361] Jessica Schroeder, Jane Hoffswell, Chia-Fang Chung, James Fogarty, Sean Munson, and Jasmine Zia. Supporting patient-provider collaboration to identify individual triggers using food and symptom journals. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, pages 1726–1739. ACM, 2017. doi: 10.1145/2998181.2998276.

[362] Jessica Schroeder, Chia-Fang Chung, Daniel A Epstein, Ravi Karkar, Adele Parsons, Natalia Murinova, James Fogarty, and Sean A Munson. Examining self-tracking by people with migraine: goals, needs, and opportunities in a chronic health condition. In *Proceedings of the 2018 designing interactive systems conference*, pages 135–148, 2018.

[363] Jessica Schroeder, Chelsey Wilkes, Kael Rowan, Arturo Toledo, Ann Paradiso, Mary Czerwinski, Gloria Mark, and Marsha M Linehan. Pocket skills: A conversational mobile web app to support dialectical behavioral therapy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2018. doi: 10.1145/3173574.3173972.

[364] Jessica Schroeder, Jina Suh, Chelsey Wilks, Mary Czerwinski, Sean A Munson, James Fogarty, and Tim Althoff. Data-driven implications for translating evidence-based psychotherapies into technology-delivered interventions. In *Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare*, pages 274–287, 2020.

[365] Stephen M Schueller. Preferences for positive psychology exercises. *The Journal of Positive Psychology*, 5(3):192–203, 2010.

[366] Thilo Schuler, Sebastian Garde, Sam Heard, and Thomas Beale. Towards auto-

matically generating graphical user interfaces from openehr archetypes. *Studies in Health Technology and Informatics*, 124:221, 2006.

[367] N. D. Schwartz, T. Hsu, and P. Cohen. Stymied in Seeking Benefits, Millions of Unemployed Go Uncounted. https://www.nytimes.com/2020/04/30/business/economy/coronavirus-unemployment-claims.html (Accessed 2021-08-17), Apr 2020.

[368] USDA Economic Research Service. 2013 rural-urban continuum codes, 2013. URL https://data.nal.usda.gov/dataset/rural-urban-continuum-codes.

[369] Adrian BR Shatte, Delyse M Hutchinson, and Samantha J Teague. Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, 49(9):1426–1448, 2019. doi: 10.1017/S0033291719000151.

[370] Joanne Shaw, Suvena Sethi, Lisa Vaccaro, Lisa Beatty, Laura Kirsten, David Kissane, Brian Kelly, Geoff Mitchell, Kerry Sherman, and Jane Turner. Is care really shared? a systematic review of collaborative care (shared care) interventions for adult cancer patients with cepression. *BMC Health Services Research*, 19(1):120, 2019. doi: 10.1186/s12913-019-3946-z.

[371] Liuhua Shi, Kyle Steenland, Haomin Li, Pengfei Liu, Yuhan Zhang, Robert H Lyles, Weeberb J Requia, Sindana D Ilango, Howard H Chang, Thomas Wingo, et al. A national cohort study (2000–2018) of long-term air pollution exposure and incident dementia in older adults in the united states. *Nature communications*, 12(1):1–9, 2021.

[372] Stuart D Sidle. Workplace stress management interventions: What works best? *Academy of Management Perspectives*, 22(3):111–112, 2008.

[373] Cynthia J Sieck, Amy Sheon, Jessica S Ancker, Jill Castek, Bill Callahan, and

Angela Siefer. Digital inclusion as a social determinant of health. *NPJ Digital Medicine*, 4(1):1–3, 2021.

[374] Gregory E Simon and Evette J Ludman. Predictors of early dropout from psychotherapy for depression in community practice. *Psychiatric Services*, 61 (7):684–689, 2010. doi: 10.1176/ps.2010.61.7.684.

[375] Gregory E Simon, Victoria Ding, Rebecca Hubbard, Paul Fishman, Evette Ludman, Leo Morales, Belinda Operskalski, and James Savarino. Early dropout from psychotherapy for depression with group-and network-model therapists. *Administration and Policy in Mental Health and Mental Health Services Research*, 39(6):440–447, 2012. doi: 10.1007/s10488-011-0364-x.

[376] Emma Simpson, Richard Brown, Elizabeth Sillence, Lynne Coventry, Karen Lloyd, Jo Gibbs, Shema Tariq, and Abigail C Durrant. Understanding the barriers and facilitators to sharing patient-generated health data using digital technology for people living with long-term health conditions: A narrative review. *Frontiers in Public Health*, 9, 2021.

[377] Sravani Singu, Arpan Acharya, Kishore Challagundla, and Siddappa N Byrareddy. Impact of social determinants of health on the emerging covid-19 pandemic in the united states. *Frontiers in public health*, page 406, 2020.

[378] Meredith M Skeels, Kenton T Unruh, Christopher Powell, and Wanda Pratt. Catalyzing social support for breast cancer patients. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 173–182. ACM, 2010. doi: 10.1145/1753326.1753353.

[379] Bruce W Smith, Jeanne Dalen, Kathryn Wiggins, Erin Tooley, Paulette Christopher, and Jennifer Bernard. The brief resilience scale: assessing the ability to bounce back. *International journal of behavioral medicine*, 15(3):194–200, 2008.

[380] Hamish R Smith. Depression in cancer patients: Pathogenesis, implications and treatment. *Oncology Letters*, 9(4):1509–1514, 2015. doi: 10.3892/ol.2015.2944.

[381] Joshua M Smyth and Kristin E Heron. Is providing mobile interventions "just-in-time" helpful? an experimental proof of concept study of just-in-time intervention for stress management. In *2016 IEEE Wireless Health (WH)*, pages 1–7. IEEE, 2016.

[382] W Söllner, A DeVries, E Steixner, P Lukas, G Sprinzl, G Rumpold, and S Maislinger. How successful are oncologists in identifying patient distress, perceived social support, and need for psychosocial counselling? *British Journal of Cancer*, 84(2):179, 2001. doi: 10.1054/bjoc.2000.1545.

[383] Ashley Southall. Why a drop in domestic violence reports might not be a good sign. https://www.nytimes.com/2020/04/17/nyregion/new-york-city-domestic-violence-coronavirus.html, journal=The New York Times (Accessed 2021-08-17), Apr 2020.

[384] JAN Spijker, Ron De Graaf, Rob V Bijl, Aartjan TF Beekman, Johan Ormel, and Willem A Nolen. Duration of major depressive episodes in the general population: Results from the netherlands mental health survey and incidence study (nemesis). *The British Journal of Psychiatry*, 181(3):208–213, 2002. doi: 10.1192/bjp.181.3.208.

[385] Robert L Spitzer, Kurt Kroenke, Janet BW Williams, and Bernd Löwe. A brief measure for assessing generalized anxiety disorder: The gad-7. *Archives of Internal Medicine*, 166(10):1092–1097, 2006. doi: 10.1001/archinte.166.10.1092.

[386] Donna Spruijt-Metz and Wendy Nilsen. Dynamic models of behavior for just-in-time adaptive interventions. *IEEE Pervasive Computing*, 13(3):13–17, 2014.

[387] Donna Spruijt-Metz, Cheng KF Wen, Gillian O'Reilly, Ming Li, Sangwon Lee, BA Emken, Urbashi Mitra, Murali Annavaram, Gisele Ragusa, and Shrikanth Narayanan. Innovations in the use of interactive technology to support weight management. *Current Obesity Reports*, 4(4):510–519, 2015. URL https://dx.doi.org/10.1007%2Fs13679-015-0183-6.

[388] Kari Staat and Milena Segatore. The phenomenon of chemo brain. *Clinical Journal of Oncology Nursing*, 9(6):713, 2005. doi: 10.1188/05.cjon.713-721.

[389] Paul Stallard, Thomas Richardson, Sophie Velleman, and Megan Attwood. Computerized cbt (think, feel, do) for depression and anxiety in children and adolescents: Outcomes and feedback from a pilot randomized controlled trial. *Behavioural and Cognitive Psychotherapy*, 39(3):273–284, 2011. doi: 10.1017/s135246581000086x.

[390] Jennifer L Steel, David A Geller, Kevin H Kim, Lisa H Butterfield, Michael Spring, Jonathan Grady, Weiing Sun, Wallis Marsh, Michael Antoni, Mary Amanda Dew, et al. Web-based collaborative care intervention to manage cancer-related symptoms in the palliative care setting. *Cancer*, 122(8):1270–1282, 2016. doi: 10.1002/cncr.29906.

[391] Zachary Steel, Claire Marnane, Changiz Iranpour, Tien Chey, John W Jackson, Vikram Patel, and Derrick Silove. The global prevalence of common mental disorders: A systematic review and meta-analysis 1980–2013. *International Journal of Epidemiology*, 43(2):476–493, 2014. doi: 10.1093/ije/dyu038.

[392] Rebecca E. Stewart, Dianne L. Chambless, and Shannon Wiltsey Stirman. Decision making and the use of evidence-based practice: Is the three-legged stool balanced? *Practice Innovations*, 3(1):56–67, 2018. ISSN 2377-889X. doi: 10.1037/pri0000063.

[393] Shannon Wiltsey Stirman, Christopher J. Miller, Katherine Toder, and Amber Calloway. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. *Implementation Science*, 8(1): 1–12, 2013. ISSN 17485908. doi: 10.1186/1748-5908-8-65.

[394] Shannon Wiltsey Stirman, Ana A. Baumann, and Christopher J. Miller. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. *Implementation Science 2019 14:1*, 14(1):1–10, jun 2019. ISSN 1748-5908. doi: 10.1186/S13012-019-0898-Y. URL https://implementationscience.biomedcentral.com/articles/10.1186/s13012-019-0898-y.

[395] Elizabeth Stratton, Nathan Jones, Susan E. Peters, John Torous, and Nicholas Glozier. Digital mHealth Interventions for Employees: Systematic Review and Meta-Analysis of Their Effects on Workplace Outcomes. *Journal of occupational and environmental medicine*, 63(8):e512–e525, 2021. ISSN 15365948. doi: 10.1097/JOM.0000000000002267.

[396] Victor J Strecher, Brenda McEvoy DeVellis, Marshall H Becker, and Irwin M Rosenstock. The role of self-efficacy in achieving health behavior change. *Health education quarterly*, 13(1):73–92, 1986.

[397] Elizabeth A Stuart. Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics*, 25(1):1, 2010.

[398] Elizabeth A Stuart, Brian K Lee, and Finbarr P Leacy. Prognostic score–based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research. *Journal of clinical epidemiology*, 66(8): S84–S90, 2013.

[399] Jina Suh, Spencer Williams, Jesse R. Fann, James Fogarty, Amy M. Bauer, and Gary Hsieh. A parallel journeys perspective on technology-supported depression care for patients with cancer. *ACM CHI Workshop on Rethinking Mental Health Resources*, 2020.

[400] Jina Suh, Spencer Williams, Jesse R Fann, James Fogarty, Amy M Bauer, and Gary Hsieh. Parallel journeys of patients with cancer and depression: challenges and opportunities for technology-enabled collaborative care. *Proceedings of the ACM on Human-computer Interaction*, 4(CSCW1):1–36, 2020.

[401] Jina Suh, Spencer Williams, Jesse R. Fann, James Fogarty, Gary Hsieh, and Amy M. Bauer. Human-centered methods to inform the design of information technologies for team-based depression care. *AMIA Clinical Informatics Conference*, 2020.

[402] Jina Suh, Eric Horvitz, Ryen W White, and Tim Althoff. Population-scale study of human needs during the covid-19 pandemic: Analysis and implications. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*, pages 4–12, 2021.

[403] Jina Suh, Eric Horvitz, Ryen W White, and Tim Althoff. Disparate impacts on online information access during the covid-19 pandemic. *Nature Communications*, 13:7094, 2022.

[404] Mason Sutherland, Mark McKenney, and Adel Elkbuli. Gun violence during covid-19 pandemic: paradoxical trends in new york city, chicago, los angeles and baltimore. *The American journal of emergency medicine*, 39:225–226, 2021.

[405] Joshua K Swift and Roger P Greenberg. Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 80(4):547, 2012. doi: 10.1037/a0028226.

[406] Michael Tanana, Kevin A Hallgren, Zac E Imel, David C Atkins, and Vivek Srikumar. A comparison of natural language processing methods for automated coding of motivational interviewing. *Journal of Substance Abuse Treatment*, 65: 43–50, 2016.

[407] Louis Tay and Ed Diener. Needs and subjective well-being around the world. *Journal of personality and social psychology*, 101(2):354, 2011.

[408] Jaime Teevan, Susan T Dumais, and Eric Horvitz. Personalizing search via automated analysis of interests and activities. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 449–456, 2005.

[409] Anja Thieme, Danielle Belgrave, and Gavin Doherty. Machine learning in mental health: A systematic review of the hci literature to support the development of effective and implementable ml systems. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(5):1–53, 2020.

[410] Anilkrishna B Thota, Theresa Ann Sipe, Guthrie J Byard, Carlos S Zometa, Robert A Hahn, Lela R McKnight-Eily, Daniel P Chapman, Ana F Abraido-Lanza, Jane L Pearson, Clinton W Anderson, et al. Collaborative care to improve the management of depressive disorders: A community guide systematic review and meta-analysis. *American Journal of Preventive Medicine*, 42(5):525–538, 2012. doi: 0.1016/j.amepre.2012.01.019.

[411] Joseph M Trombello, Charles South, Audrey Cecil, Katherine E Sánchez, Alma Christina Sánchez, Sara Levinson Eidelman, Taryn L Mayes, Farra Kahalnik, Corey Tovian, and Beth D Kennard. Efficacy of a behavioral activation teletherapy intervention to treat depression and anxiety in primary care vitalsign6 program. *The primary care companion for CNS disorders*, 19(5):26691, 2017.

[412] Emily Troshynski, Charlotte Lee, and Paul Dourish. Accountabilities of presence: reframing location-based systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 487–496, 2008.

[413] Allison S Troy and Iris B Mauss. Resilience in the face of stress: Emotion regulation as a protective factor. *Resilience and mental health: Challenges across the lifespan*, 1(2):30–44, 2011.

[414] Annie Tubadji, Frederic Boy, and D Webber. Narrative economics, public policy and mental health. *Center for Economic Policy Research*, 20:109–131, 2020.

[415] Jürgen Unützer, Henry Harbin, Michael Schoenbaum, and Benjamin Druss. The collaborative care model: An approach for integrating physical and mental health care in medicaid health homes. *HEALTH HOME, Information Resource Center*, pages 1–13, 2013. doi: 10.1176/appi.ps.57.1.37.

[416] U.S. Census Bureau. State Area Measurements and Internal Point Coordinates. https://www.census.gov/geographies/reference-files/2010/geo/state-area.html (Accessed 2020-06-08), .

[417] U.S. Census Bureau. ZIP Code Tabulation Areas. https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2019.html (Accessed 2020-06-08), .

[418] U.S. Census Bureau. American Community Survey 5-year estimates (2014-2018). https://censusreporter.org (Accessed 2020-06-08), .

[419] U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. Social Determinants of Health - Healthy People 2030. https://health.gov/healthypeople/objectives-and-data/social-determinants-health (Accessed 2021-08-17).

[420] Kim Usher, Navjot Bhullar, and Debra Jackson. Life in the pandemic: Social isolation and mental health. *Journal of Clinical Nursing*, 2020.

[421] Margot Van Der Doef and Stan Maes. The Job Demand-Control(-Support) model and psychological well-being: A review of 20 years of empirical research. *Work and Stress*, 13(2):87–114, 1999. ISSN 02678373. doi: 10.1080/026783799296084.

[422] Alexander JAM Van Deursen and Ellen J Helsper. The third-level digital divide: Who benefits most from being online? In *Communication and information technologies annual*. Emerald Group Publishing Limited, 2015.

[423] Alexander JAM Van Deursen and Jan AGM Van Dijk. The first-level digital divide shifts from inequalities in physical access to inequalities in material access. *New media & society*, 21(2):354–375, 2019.

[424] Frank J van Lenthe, Tessa Jansen, and Carlijn BM Kamphuis. Understanding socio-economic inequalities in food choice behaviour: can maslow's pyramid help? *British Journal of Nutrition*, 113(7):1139–1147, 2015.

[425] ER Vanderlip, J Rundell, M Avery, C Alter, C Engel, J Fortney, and M Williams. Dissemination of integrated care within adult primary care settings: The collaborative care model. *Washington DC: American Psychiatric Association and Academy of Psychosomatic Medicine*, 2016.

[426] Viswanath Venkatesh, Michael G Morris, Gordon B Davis, and Fred D Davis. User acceptance of information technology: Toward a unified view. *MIS quarterly*, pages 425–478, 2003.

[427] Tanja GM Vrijkotte, Lorenz JP Van Doornen, and Eco JC De Geus. Effects of work stress on ambulatory blood pressure, heart rate, and heart rate variability. *Hypertension*, 35(4):880–886, 2000.

[428] Amy W Wagner, Matthew Jakupcak, Halina M Kowalski, Joyce N Bittinger, and Shahrokh Golshan. Behavioral activation as a treatment for posttraumatic stress disorder among returning veterans: A randomized trial. *Psychiatric services*, 70 (10):867–873, 2019.

[429] Mahmoud A. Wahba and Lawrence G. Bridwell. Maslow reconsidered: A review of research on the need hierarchy theory. *Organ. Behav. Hum. Decis. Process*, 15 (2), 1976.

[430] Jane Walker, Christian Holm Hansen, Paul Martin, Stefan Symeonides, Ravi Ramessur, Gordon Murray, and Michael Sharpe. Prevalence, associations, and adequacy of treatment of major depression in patients with cancer: A cross-sectional analysis of routinely collected clinical data. *The Lancet Psychiatry*, 1 (5):343–350, 2014. doi: 10.1016/s2215-0366(14)70313-x.

[431] Liyuan Wang and Lynn Carol Miller. Just-in-the-moment adaptive interventions (jitai): A meta-analytical review. *Health Communication*, 35(12):1531–1544, 2020.

[432] Manjie Wang and Kimberly J Saudino. Emotion regulation and stress. *Journal of Adult Development*, 18(2):95–103, 2011.

[433] Rui Wang, Andrew T Campbell, and Xia Zhou. Using opportunistic face logging from smartphone to infer mental health: Challenges and future directions. In *Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pages 683–692. ACM, 2015. doi: 10.1145/2800835.2804391.

[434] Rui Wang, Min SH Aung, Saeed Abdullah, Rachel Brian, Andrew T Campbell, Tanzeem Choudhury, Marta Hauser, John Kane, Michael Merrill, Emily A Scherer, et al. Crosscheck: Toward passive sensing and detection of mental

health changes in people with schizophrenia. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 886–897. ACM, 2016. doi: 10.1145/2971648.2971740.

[435] Paul Waraich, Elliot M Goldner, Julian M Somers, and Lorena Hsu. Prevalence and incidence studies of mood disorders: A systematic review of the literature. *The Canadian Journal of Psychiatry*, 49(2):124–138, 2004. doi: 10.1177/070674370404900208.

[436] Elizabeth Ward, Recinda L Sherman, S Jane Henley, Ahmedin Jemal, David A Siegel, Eric J Feuer, Albert U Firth, Betsy A Kohler, Susan Scott, Jiemin Ma, et al. Annual report to the nation on the status of cancer, 1999–2015, featuring cancer in men and women ages 20–49. *JNCI: Journal of the National Cancer Institute*, 2019. doi: 10.1093/jnci/djz106.

[437] Ingmar Weber, Venkata Rama Kiran Garimella, and Erik Borra. Mining web query logs to analyze political issues. In *Proceedings of the 4th annual acm web science conference*, pages 330–334, 2012.

[438] Rina R Wehbe, Diane K Watson, Gustavo F Tondello, Marim Ganaba, Melissa Stocco, Alvin Lee, and Lennart E Nacke. Above water: An educational game for anxiety. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, pages 79–84. ACM, 2016. doi: 10.1145/2968120.2971804.

[439] Bryan J Weiner, Cara C Lewis, Cameo Stanick, Byron J Powell, Caitlin N Dorsey, Alecia S Clary, Marcella H Boynton, and Heather Halko. Psychometric assessment of three newly developed implementation outcome measures. *Implementation Science*, 12(1):1–12, 2017.

[440] Gregory A. Wellenius, Swapnil Vispute, Valeria Espinosa, Alex Fabrikant,

Thomas C. Tsai, Jonathan Hennessy, Brian Williams, Krishina Gadepalli, Adam Boulange, Adam Pearce, et al. Impacts of state-level policies on social distancing in the united states using aggregated mobility data during the covid-19 pandemic, 2020.

[441] Amy Wenzel. Basic strategies of cognitive behavioral therapy. *Psychiatric Clinics of North America*, 2017.

[442] Robert West, Ryen W White, and Eric Horvitz. From cookies to cooks: Insights on dietary patterns via analysis of web usage logs. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1399–1410, 2013.

[443] Robert West, Ryen W White, and Eric Horvitz. Here and there: Goals, activities, and predictions about location from geotagged queries. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pages 817–820, 2013.

[444] Mina Westman and Dalia Etzion. The impact of vacation and job stress on burnout and absenteeism. *Psychology & Health*, 16(5):595–606, 2001.

[445] Ryen W White and Eric Horvitz. Evaluation of the feasibility of screening patients for early signs of lung carcinoma in web search logs. *JAMA oncology*, 3(3):398–401, 2017.

[446] Ryen W White and Dan Morris. Investigating the querying and browsing behavior of advanced search engine users. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 255–262, 2007.

[447] Ryen W White, Peter Bailey, and Liwei Chen. Predicting user interests from contextual information. In *Proceedings of the 32nd international ACM SIGIR*

conference on Research and development in information retrieval, pages 363–370, 2009.

[448] Ryen W White, Susan T Dumais, and Jaime Teevan. Characterizing the influence of domain expertise on web search behavior. In *Proceedings of the second ACM international conference on web search and data mining*, pages 132–141, 2009.

[449] David R Williams and Chiquita Collins. Racial residential segregation: a fundamental cause of racial disparities in health. *Public health reports*, 2016.

[450] David R. Williams, Manuela Costa, Adebola O Odunlami, and Selina A. Mohammed. Moving upstream: how interventions that address the social determinants of health can improve health and reduce disparities. *Journal of public health management and practice : JPHMP*, 14:S8–17, 2008.

[451] David R Williams, Selina A Mohammed, Jacinta Leavell, and Chiquita Collins. Race, socioeconomic status and health: Complexities, ongoing challenges and research opportunities. *Annals of the New York Academy of Sciences*, 1186:69, 2010.

[452] Coady Wing, Kosali Simon, and Ricardo A Bello-Gomez. Designing difference in difference studies: best practices for public health policy research. *Annu Rev Public Health*, 39(1):453–469, 2018.

[453] Emily Wood, Sally Ohlsen, and Thomas Ricketts. What are the barriers and facilitators to implementing collaborative care for depression? a systematic review. *Journal of Affective Disorders*, 214:26–43, 2017. doi: 10.1016/j.jad.2017.02.028.

[454] Susan S Woods, Neil C Evans, and Kathleen L Frisbee. Integrating patient voices into health information for self-care and patient-clinician partnerships: Veterans

affairs design recommendations for patient-generated data applications. *Journal of the American Medical Informatics Association*, 23(3):491–495, 2016.

[455] World Health Organization. A conceptual framework for action on the social determinants of health. *World Health Organization*, 2010.

[456] Maja Wrzesien, Mariano Alcañiz, Cristina Botella, Jean-Marie Burkhardt, Juana Breton Lopez, and Alejandro Rodriguez Ortega. A pilot evaluation of a therapeutic game applied to small animal phobia treatment. In *International Conference on Serious Games Development and Applications*, pages 10–20. Springer, 2014. doi: 10.1007/978-3-319-11623-5_2.

[457] Shinyi Wu, Kathleen Ell, Haomiao Jin, Irene Vidyanti, Chih-Ping Chou, Pey-Jiuan Lee, Sandra Gross-Schulman, Laura Myerchin Sklaroff, David Belson, Arthur M Nezu, et al. Comparative effectiveness of a technology-facilitated depression care management model in safety-net primary care patients with type 2 diabetes: 6-month outcomes of a large clinical trial. *Journal of Medical Internet Research*, 20(4):e147, 2018. doi: 10.2196/jmir.7692.

[458] Clyde W Yancy. COVID-19 and african americans. *Jama*, 323(19):1891–1892, 2020.

[459] Huahai Yang and Yunyao Li. Identifying user needs from social media. *IBM Research Division, San Jose*, 2013.

[460] Lucy Yardley, Bonnie J Spring, Heleen Riper, Leanne G Morrison, David H Crane, Kristina Curtis, Gina C Merchant, Felix Naughton, and Ann Blandford. Understanding and promoting effective engagement with digital behavior change interventions. *American journal of preventive medicine*, 51(5):833–842, 2016.

[461] Andrew Yates, Arman Cohan, and Nazli Goharian. Depression and self-harm risk assessment in online forums. *arXiv preprint arXiv:1709.01848*, 2017.

[462] Jiali Ye, George Rust, Yvonne Fry-Johnson, and Harry Strothers. E-mail in patient–provider communication: A systematic review. *Patient Education and Counseling*, 80(2):266–273, 2010. doi: 10.1016/j.pec.2009.09.038.

[463] Baligh R Yehia, Angela Winegar, Richard Fogel, Mohamad Fakih, Allison Ottenbacher, Christine Jesser, Angelo Bufalino, Ren-Huai Huang, and Joseph Cacchione. Association of race with mortality among patients hospitalized with coronavirus disease 2019 (COVID-19) at 92 US hospitals. *JAMA network open*, 3 (8):e2018039–e2018039, 2020.

[464] Matthew Yglesias. Reopening schools safely is going to take much more federal leadership. https://www.vox.com/2020/7/8/21314563/school-reopening-testing-money (Accessed 2021-08-17), Jul 2020.

[465] Zearn.org. Press Release: Zearn Provides Real-Time Snapshot on the State of U.S. Math Education Through New OI Economic Tracker by Opportunity Insights. https://about.zearn.org/press-releases/zearn-provides-real-time-snapshot-on-the-state-of-u-s-math-education-through-new-oi-economic-tracker-by-opportunity-insights (Accessed 2021-08-17), 2021.

[466] Brad Zebrack, Karen Kayser, Lynne Padgett, Laura Sundstrom, Chad Jobin, Krista Nelson, and Iris C Fineberg. Institutional capacity to provide psychosocial oncology support services: A report from the association of oncology social work. *Cancer*, 122(12):1937–1945, 2016. doi: 10.1002/cncr.30016.

[467] Min-Ling Zhang and Zhi-Hua Zhou. A review on multi-label learning algorithms. *IEEE transactions on knowledge and data engineering*, 26(8), 2013.

Appendix A

MEASURING HUMAN NEEDS THROUGH SEARCH DATA

A.1 ANALYZING NATIONAL REPRESENTATION

To understand how much of the US population is represented by the collected data, we obtained demographics data from the Census Reporter API [418] for ZIP codes represented in our dataset. Census Reporter API provides demographics data for 32,989 US ZIP codes, and not all of the ZIP codes in our dataset have available demographics information through this service. Between our dataset and the demographics data, we have 96.4% overlap of ZIP codes, representing 97.5% of total queries in our dataset. Table A.1 summarizes the median of the select 11 demographic variables for ZIP codes in our dataset in comparison to all available ZIP codes in Census Reporter. Given 96.4% overlap, we find that the ZIP codes in our dataset closely mirrors the US population. Although query volumes are not uniformly distributed across these ZIP codes, the vast majority of US ZIP codes is included in our dataset.

A.2 ANALYZING SELECTION BIAS

We sought to understand potential biases in socioeconomic circumstances that would influence the usage of Bing search engine. We leveraged deidentified client id as a proxy for a unique user to estimate the 'client rate', or how much of the population in a ZIP code is using the Bing search engine. We examined the correlation between the client rate and various demographic factors at the ZIP code level (e.g., income, race, age, gender, education, housing, internet access). Although unsurprisingly the factors

	Our Dataset	Census
Population	3,072	2,775
Median Income	$54,231	$54,048
Median Age	41.8	41.9
% Race White	0.87	0.88
% Male	0.5	0.5
% HS Grad or Higher	0.9	0.9
Gini Index	0.42	0.42
% Below Poverty Lvl.	0.12	0.12
% Housing Owned	0.76	0.76
% Has Internet	0.77	0.76

Table A.1: Distribution of demographic variables (median) for US ZIP codes in our dataset compared to all of the available US ZIP codes in the census data.

Demographic Variable	Corr Coeff
% Housing Owned	-0.058^{***}
% Female	-0.029^{***}
% Race White	-0.024^{***}
Median Income	0.021^{***}
% Has Internet	0.019^{***}
% HS Grad or Higher	0.018^{***}
Median Age	-0.010
% Below Poverty Level	-0.007
$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$	

Table A.2: Pearson correlation between Bing client rate and demographic variables for each ZIP code.

describe some of the variance, none were correlated more strongly than r=-0.058 (% Housing Owned). Table A.2 summarizes Pearson correlation statistics. These results suggest that our dataset is not strongly biased towards any single demographic.

A.3 RELIABILITY OF SEARCH INTERACTION TRENDS

Although our analysis relies solely on Bing search data, many Americans use other search engines such as Google. Therefore, we compared search trends for Bing with data available via the Google Trends API[1] for the same time period and a subset

---

[1] https://trends.google.com/

Need	Keyword	Corr Coeff
SA	recipe	0.960^{***}
SA	netflix	0.935^{***}
SA	"online games with friends"	0.896^{***}
Cog	"online learning"	0.973^{***}
L&B	"online dating"	0.448^{***}
Safe	unemployment	0.977^{***}
Safe	"hand sanitizer"	0.966^{***}
Safe	tax	0.913^{***}
Safe	gun	0.764^{***}
Phys	"grocery delivery"	0.980^{***}
Phys	coronavirus	0.964^{***}
Phys	"food stamp"	0.963^{***}
Phys	health	0.888^{***}

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table A.3: Pearson correlation between Bing search trend and Google search trend for each keyword. Quotation marks indicate that the entire string is matched.

of keywords. We chose to use select keywords from each need category, rather than applying our needs aggregation pipeline (described in Sec. 3.1.2), because the Google Trends public API does not support regular expressions or access to the click interactions. We chose 9 keywords that were representative of their respective need subcategory (i.e., 'online learning' captures *online education queries*), had a significantly larger query volume compared to other keywords in the need categories (i.e., 'hand sanitizer' had more query volume than 'mask'), or may have a seasonal effect (i.e., query volume for 'tax' could depend on the tax season). We conducted a correlation analysis on a moving average of a full week to account for timezone differences between the two data sources. Visual inspection of both Google and Bing trends confirm that search patterns across these two search engines are remarkably similar (Fig. A.1), and Pearson correlation coefficients are very high with a median of 0.96 (min=0.45, max=0.98, all p<0.001). Table A.3 summarizes Pearson correlation outputs. These results imply that our findings are not simply an artifact of using one search engine over another.

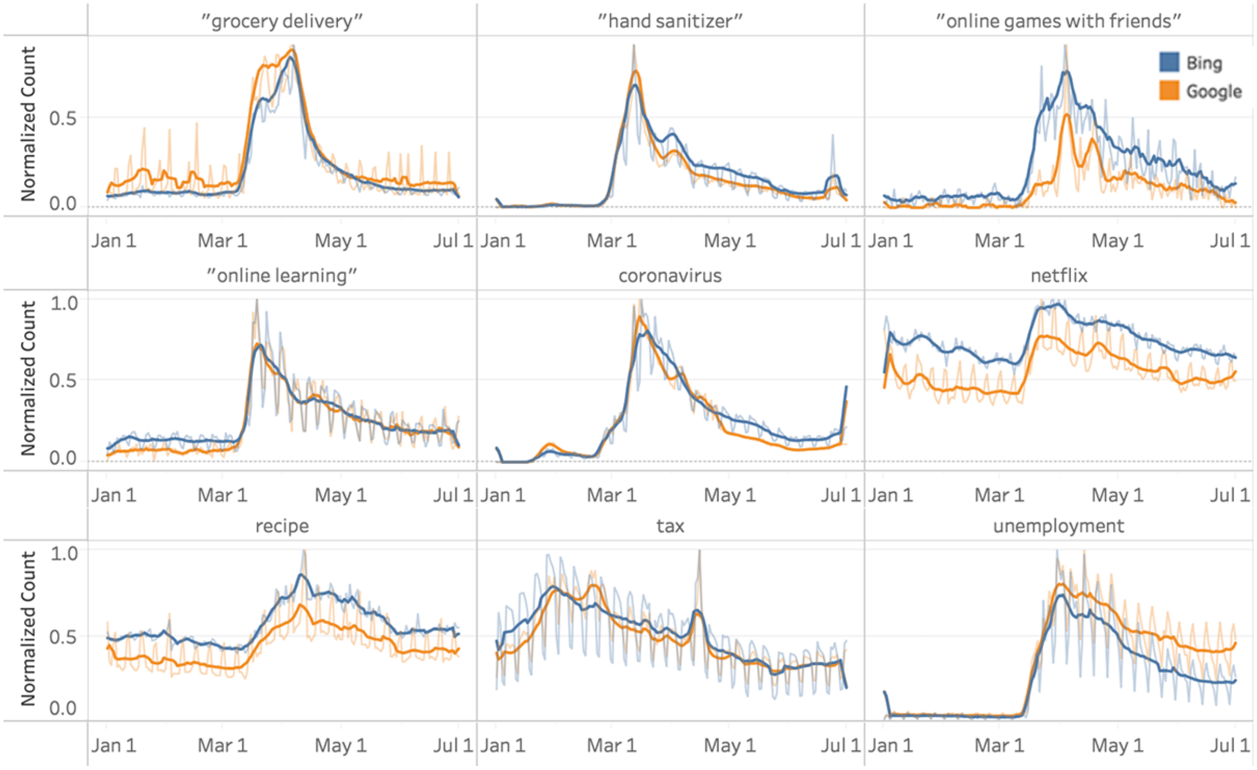


Figure A.1: Comparison of Bing and Google trends data on select keywords. The query counts are normalized where 1.0 denotes the maximum query count in each keyword trend. Bing trends (blue) closely follow Google trends (orange) with a median correlation of $r = 0.96$.

A.4 HUMAN NEEDS CATEGORIES AND DETECTION

We draw inspiration from Maslow's hierarchy of fundamental human needs [264, 265] to tag each search interaction. Of the eight top-level human needs from Maslow's expanded hierarchy of needs, we omit *Esteem*, *Aesthetics*, and *Transcendence* because they are difficult to operationalize from observational data alone. We categorize search behaviors into one or more of five broad categories of needs. *Safety* and *Physiological* are considered as 'basic' needs. *Love and Belonging*, *Cognitive*, and *Self Actualization* are often considered to be 'psychological' needs, and *Cognitive* and *Self Actualization* are considered as 'growth' needs.

We obtain further granularity of the 5 human need categories by decomposing each need category into many sub-need categories. This extra level of granularity has advantages in teasing apart the specific aspects of the need (e.g., shelter vs. finances in *Safety* needs) and in debugging and refining the detection logic. Table A.4 enumerates 79 sub-needs and their examples.

For each need subcategory, we had multiple regular expressions applied to either the query string, the clicked URL, or both. This is noted in Table A.4 under the 'Logic' column. 'Keyword and domain (KD)' logic indicates that there is one regular expression for matching query strings and another regular expression for matching clicked URLs. Both must be matched for the search interaction to be categorized as that need. 'Queries (Q)' logic indicates that there is one regular expression for query strings alone regardless of the clicked URLs. 'Domains (D)' logic indicates that there is one regular expression for clicked URLs regardless of the query strings. A full table of regular expressions used in the data collection is provided in a separate supplementary file and can be cross-referenced through the 'Need Id' column[2].

There were several steps in arriving at these subcategories and regular expressions. We first identified a set of e-commerce websites (1) from market research blog posts and reports that were publicly available online and (2) from enumerating top 100 URL hosts from a sample of search interactions that were tagged as purchase or commerce related using built-in classifiers in Bing. We combined these two data sources to manually curate a list of 61 e-commerce domains.

Next, we collected another sample of query strings that led to subsequent clicks to these e-commerce domains as well as page snippets that were displayed on the search result page. Using the Latent Dirichlet Allocation module in gensim, a popular topic modeling tool in python, we trained an unsupervised topic model on the query strings and the associated page snippets for 100, 500, and 1000 topics on 100,000 unigrams.

---

[2] https://github.com/jinasuh/pandemic_needs

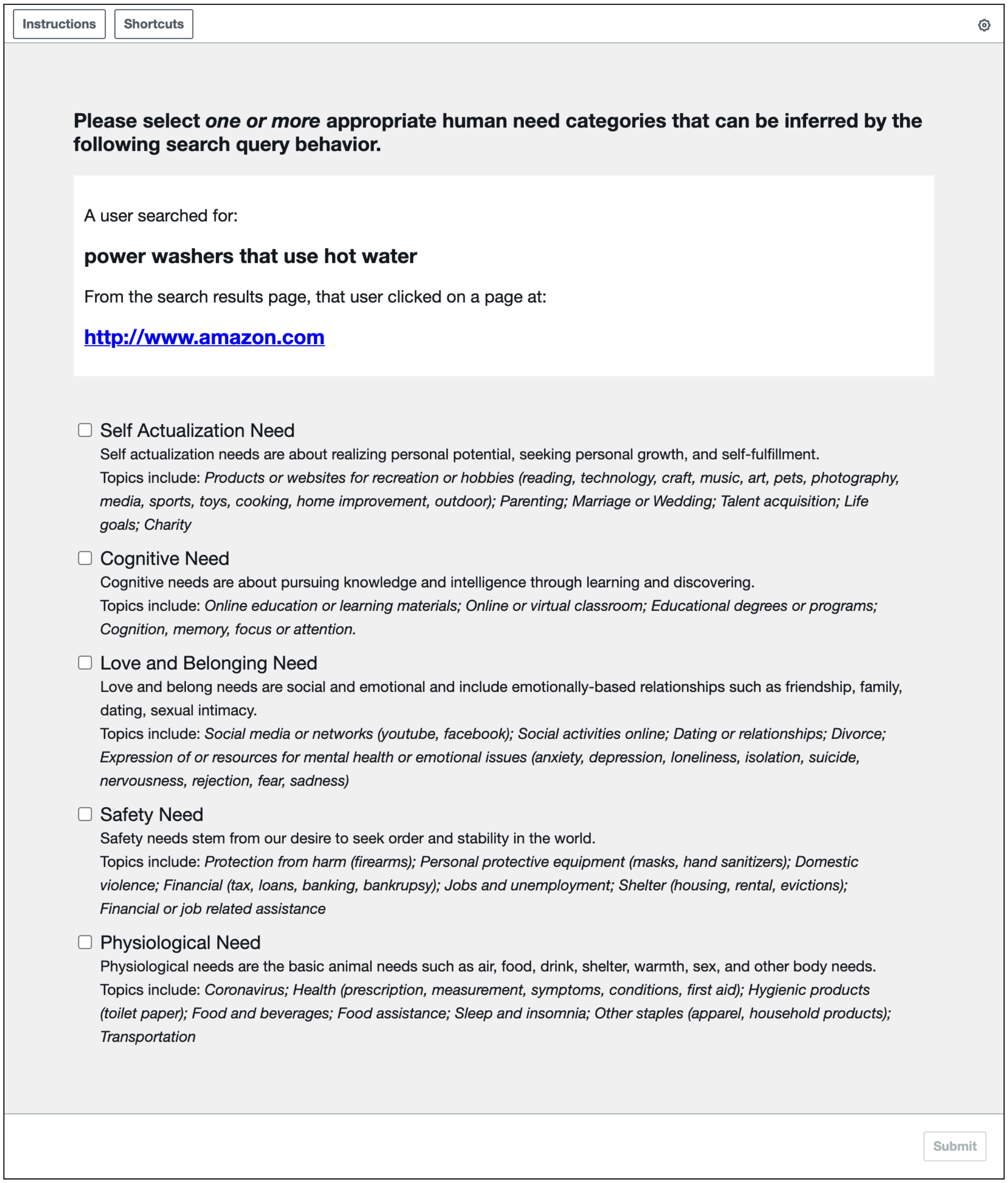
We manually inspected the keywords in the topics to curate 22 topics of over 800 keywords. We also extracted frequent bigrams from the same dataset to expand our keywords if existing unigrams were too ambiguous. These topics represented purchase categories because they were collected from search interactions with subsequent clicks to e-commerce websites.

Our next step was to categorize these topics into five main need categories. We used a hybrid card sorting method [8] with 5 participants who merged, split, edited, or created these topics into need subcategories. Then these need subcategories were then placed into the 5 main need categories.

Using the outputs from the card sorting activity as a basis, we further brainstormed subcategories beyond purchase categories. Several researchers independently brainstormed and categorized subcategories which were combined and resolved collaboratively through consensus meetings. We referred to the definitions presented in the theories of human needs to resolve any remaining disagreements [264].

Once the need categories and subcategories were identified, we leveraged keywords and domains from manually curated topics and card sort outputs. We also obtained a list of topical domains from publicly available blog posts and articles summarizing recommended websites (e.g., search for best parenting website) or manually curated a list of domains (e.g., search for unemployment benefit page for each state). Again, we independently brainstormed and curated lists of keywords, keyword patterns, and domains for each subcategory and collaboratively iterated on the lists through consensus meetings. These keywords, keyword patterns, and domain URLs were combined into many regular expressions that we used to tag each search interaction.

We validated our detection logic by collecting human labels from Amazon's Mechanical Turk, and our evaluation set achieved a precision of 97.2%, as described in detail in Section 3.1.2. A screenshot of an example task seen by a crowd worker is shown in Fig. A.2, and the full set of detailed task instructions can be found in Appendix A.5.



Please select *one or more* appropriate human need categories that can be inferred by the following search query behavior.

A user searched for:

power washers that use hot water

From the search results page, that user clicked on a page at:

http://www.amazon.com

☐ Self Actualization Need
Self actualization needs are about realizing personal potential, seeking personal growth, and self-fulfillment.
Topics include: *Products or websites for recreation or hobbies (reading, technology, craft, music, art, pets, photography, media, sports, toys, cooking, home improvement, outdoor); Parenting; Marriage or Wedding; Talent acquisition; Life goals; Charity*

☐ Cognitive Need
Cognitive needs are about pursuing knowledge and intelligence through learning and discovering.
Topics include: *Online education or learning materials; Online or virtual classroom; Educational degrees or programs; Cognition, memory, focus or attention.*

☐ Love and Belonging Need
Love and belong needs are social and emotional and include emotionally-based relationships such as friendship, family, dating, sexual intimacy.
Topics include: *Social media or networks (youtube, facebook); Social activities online; Dating or relationships; Divorce; Expression of or resources for mental health or emotional issues (anxiety, depression, loneliness, isolation, suicide, nervousness, rejection, fear, sadness)*

☐ Safety Need
Safety needs stem from our desire to seek order and stability in the world.
Topics include: *Protection from harm (firearms); Personal protective equipment (masks, hand sanitizers); Domestic violence; Financial (tax, loans, banking, bankrupsy); Jobs and unemployment; Shelter (housing, rental, evictions); Financial or job related assistance*

☐ Physiological Need
Physiological needs are the basic animal needs such as air, food, drink, shelter, warmth, sex, and other body needs.
Topics include: *Coronavirus; Health (prescription, measurement, symptoms, conditions, first aid); Hygienic products (toilet paper); Food and beverages; Food assistance; Sleep and insomnia; Other staples (apparel, household products); Transportation*

Figure A.2: Example labeling task for Amazon's Mechanical Turk crowd worker. A query string and/or a clicked domain URL is displayed with multi-select options for five need categories. Each option includes a brief description and a set of example topics for the category.

Need Category	Need Id	Subcategory	Logic	Example
SelfActualization	SA1	Audio/video related purchase	KD	home theater + e-commerce url click
	SA2	Books and reading related purchase	KD	comics + e-commerce url click
	SA3	Charity related queries	Q	charity; donations
	SA4	Cooking related purchase	KD	cookbooks + e-commerce url click
	SA5	Cooking related queries	Q	how to saute onions
	SA6	Cooking site visits	D	click on foodnetwork.com
	SA7	Crafts instruction queries	Q	how to embroider
	SA8	Crafts related purchase	KD	scrapbooking + e-commerce url click
	SA9	DIY site visits	D	click on instructables.com
	SA10	Home improvement purchase	KD	circular saw + e-commerce url click
	SA11	Home improvement related queries	Q	remodeling; wood working
	SA12	Home improvement site visits	D	click on homedepot.com
	SA13	Life goal related queries	Q	goals for living; personal goals
	SA14	Media related purchase	KD	dvd player + e-commerce url click
	SA15	Music instruction queries	Q	learn to play guitar; how to play a drum
	SA16	Musical instrument related purchase	KD	piano + e-commerce url click
	SA17	Outdoor related purchase	KD	hiking shoes + e-commerce url click
	SA18	Outdoor related queries	Q	best national parks
	SA19	Outdoor site visits	D	click on rei.com
	SA20	Parenting site visits	D	click on parents.com
	SA21	Pet related purchase	KD	ferret + e-commerce url click
	SA22	Pet related queries	Q	care for frogs
	SA23	Pet related site visits	D	click on petco.com
	SA24	Photography related purchase	KD	photography + e-commerce url click
	SA25	Photography related queries	Q	how to take pictures of a bird
	SA26	Sports related purchase	KD	elliptical trainer + e-commerce url click
	SA27	Streaming media site visits	D	click on netflix.com
	SA28	Technology related purchase	KD	ethernet + e-commerce url click
	SA29	Toys and gaming purchase	KD	xbox + e-commerce url click
	SA30	Toys and gaming site visits	D	click on twitch.tv
	SA31	Wedding related purchase	KD	wedding decorations + e-commerce url click
	SA32	Wedding site visits	D	click on theknot.com
Cognitive	C1	Cognition related queries	Q	improve memory; cant pay attention
	C2	Educational degree related queries	Q	degree program; online diploma
	C3	Educational site visits	D	click on coursera.org

	C4	Online education queries	Q	learn remotely; lesson plans
LoveBelonging	LB1	Dating related queries	Q	online dating; relationship advice
	LB2	Dating site visits	D	click on tinder.com
	LB3	Divorce related queries	Q	divorce lawyer; file for divorce
	LB4	Mental health experiential queries	Q	im alone; i feel depressed
	LB5	Mental health resource site visits	D	click on talkspace.com
	LB6	Online social activities queries	Q	online activities with family
	LB7	Social network site visits	D	click on facebook.com
	LB8	Social technology uses	KD	families + click on whatsapp.com
Safety	S1	Bank related queries	Q	banks; banking
	S2	Bankruptcy related queries	Q	foreclosure;bankruptcy
	S3	COVID-19 protection purchase	KD	disposable masks; sanitizers
	S4	Domestic violence queries	Q	domestic assult; abusive relationship
	S5	Eviction related queries	Q	evicted; rent moratorium
	S6	Financial loan related queries	Q	borrower; mortgage
	S7	Firearm purchase	KD	glock holster + e-commerce url click
	S8	Housing related queries	Q	best neighborhoods; housing safety
	S9	Job search related queries	Q	job application; resume
	S10	Job search site visits	D	click on indeed.com
	S11	Rental related queries	Q	rental apartments; houses for rent
	S12	State unemployment site visits	D	click on www.michigan.gov/uia
	S13	Stimulus related queries	Q	relief fund; loan forgiveness
	S14	Tax related queries	Q	irs; tax
	S15	Unemployment related queries	Q	im unemployed; jobless benefits
Physiological	P1	Apparel purchase	KD	athletic wear + e-commerce url click
	P2	Automobile related purchase	KD	motorcycle + e-commerce url click
	P3	Beverage purchase	KD	coffee + e-commerce url click
	P4	Cookware purchase	KD	cookie sheet + e-commerce url click
	P5	Food and beverage related queries	KD	applesauce; noodle soup
	P6	Food assistance related queries	Q	food stamps; snap program
	P7	Food delivery related queries	Q	grocery delivery; deliver meal
	P8	Food delivery site visits	D	click on instacart.com
	P9	Grocery related queries	Q	grocery; groceries
	P10	Grocery site visits	D	click on albertsons.com
	P11	Health condition related queries	Q	arthritis; diabetes
	P12	Health first aid purchase	KD	bandaid; wound closure

	P13	Health measurement equipment purchase	KD	oximeter + e-commerce site visits
	P14	Health symptom related queries	Q	i lost hearing; blurred vision
	P15	Household good purchase	KD	cleaning supplies + e-commerce url click
	P16	Insomnia related queries	Q	I cant sleep; help falling asleep
	P17	Prescription related queries	Q	medication interactions; pharmacies
	P18	Prescription site visits	D	click on rxlist.com
	P19	Sleep aid purchase	KD	sleep supplement; melatonin
	P20	Toilet paper purchase	KD	toilet paper; cottenelle

Table A.4: Human need categories and subcategories with examples. Under the Logic column, ‘KD’ refers to matching both query string and clicked URL, ‘Q’ refers to matching query string only, and ‘D’ refers to matching clicked URL only. Need Id is provided to cross-reference with a full regular expression table provided in a separate file.

A.5 HUMAN NEED DETECTION INSTRUCTIONS FOR MTURK WORKERS

A search query can be an expression of an underlying human need.

For example, if I am running a fever and want to purchase a fever reducer online (physiological need), my query might be "tylenol" with a subsequent click to amazon.com in the search results to purchase the item. If I am pursuing baking as a hobby (self actualization need), my query might be "how to bake a perfect sourdough", and the subsequent clicks do not matter much to understand the intent for that query. Or if I am visiting coursera.org, it might indicate that I want to learn a new topic (cognitive need).

In this task, please read the search query and the subsequent website visit (if available) to determine which of the following human needs category it belongs to. Search queries should seek information, express interest, intend to purchase or obtain goods that could satisfy these needs.

Self Actualization Need Self actualization needs are about realizing personal potential, seeking personal growth, and self-fulfillment. Some of the topics to consider are:

- Recreation or hobbies such as reading, technology, craft, music, art, pets, photography, media, sports, toys, cooking, home improvement, outdoor
- Parenting and child rearing
- Marriage or wedding
- Talent acquisition
- Life or personal goals
- Charity, donations, volunteering

Example 1:

Search Query: how to photograph birds in flight

Example 2:

Search Query: wedding invitation design

Visited Url: theknot.com

Cognitive Need Cognitive needs are about pursuing knowledge and intelligence through learning and discovering. Some of the topics to consider are:

- Education or learning materials
- Online or virtual classroom
- Educational degrees or programs
- Cognition, memory, focus or attention

Example 1:

Search Query: lesson plans for 4th grade math

Example 2:

Search Query: python data science

Visited Url: coursera.org

Love and Belonging Need Love and belong needs are social and emotional and include emotionally-based relationships such as friendship, family, dating, sexual intimacy. Some of the topics to consider are:

- Expression of or resources for mental health or emotional issues such as anxiety, depression, loneliness, isolation, suicide, nervousness, rejection, fear or sadness
- Social media, social network and relevant technologies
- Social activities online or offline
- Search for relationships, significant others, dating
- Negative relationships such as divorce or breakup

Example 1:

Search Query: online therapist for anxiety

Visited Url: talkspace.com

Example 2:

Search Query: long distance relationships

Safety Need Safety needs stem from our desire to seek order and stability in the world. These needs provide protection from elements through shelter and security of body, jobs, food, resources, family, or health. Some of the topics to consider are:

- Protection from harm through equipments such as firearms or security systems
- Personal protective equipment such as masks or sanitizers
- Lack of protection such as domestic violence
- Financial related such as tax, loans, banking, or bankrupsy
- Jobs and unemployment
- Shelter related such as housing, rental, evictions
- Assistance for jobs or finances

Example 1:

Search Query: n-95 masks

Visited Url: amazon.com

Example 2:

Search Query: unemployment benefits

Physiological Need Physiological needs are the basic animal needs such as air, food, drink, shelter, warmth, sex, and other body needs. Some of the topics to consider are:

- Health related such as prescription, measurement, symptoms, conditions, or first aid
- Hygienic products such as toilet paper
- Basic staples such as food, beverages, apparel, household products
- Services that provice basic staples like grocery delivery or online grocers
- Sleep and insomnia
- Assistance for food
- Transportation and mobility

Example 1:

Search Query: Symptoms of flu

Example 2:

Search Query: whole foods

Visited Url: wholefoodsmarket.com

A.6 SIGNIFICANT CHANGES IN HUMAN NEEDS

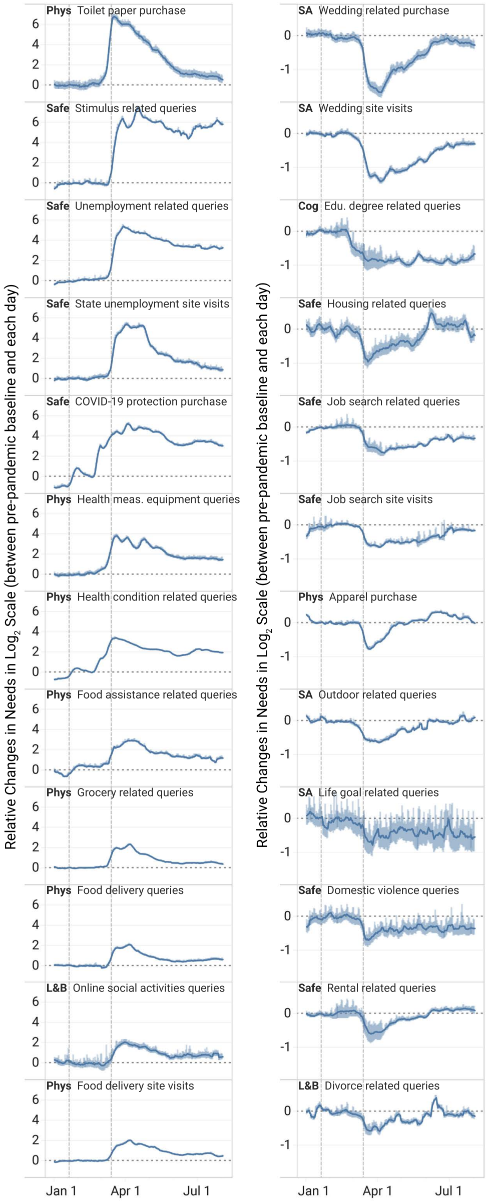


Figure A.3: Daily relative changes in needs for top 12 need subcategories with the largest increase (left) and top 12 need subcategories with the largest decrease (right). Vertical bars denote the first reported US COVID case (Jan 20) and the US national emergency declaration (Mar 13).

Appendix B

QUANTIFYING DIGITAL DISPARITIES THROUGH SEARCH DATA

As described in our Methods, we compute 95% confidence intervals through bootstrapping with 500 iterations. These confidence intervals are computed when estimating the effect size (i.e., the difference between matched groups). Therefore, they are shown as error bars in figures wherever we conduct a comparison across groups – such as all of the even numbered Supplementary Figures B.14 to B.26.

B.1 SDOH FACTORS USED IN ANALYSIS

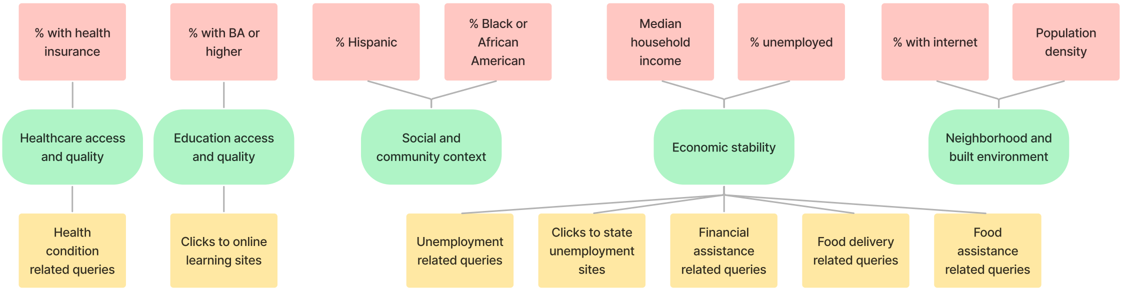


Figure B.1: Our choice of census variables (red) and search categories (yellow) are inspired by the Social Determinants of Health framework (green).

Table B.1: Means and medians of eight SDoH factors across 25,801 ZIP codes in our dataset. Population density is computed from the total U.S. population [418] divided by the total U.S. land area [416].

SDoH Factor	ZIP Code Mean	ZIP Code Median	U.S. Mean	ACS Table
Median Household Income	$60,819.849	$55,224	$65,712	B19013
% Unemployed	3.4%	3.0%	2.9%	B23025
% Black or African American Alone	8.3%	1.6%	12.8%	B02001
% Hispanic or Latino	10.1%	4.0%	18.4%	B03003
% with Bachelor or Higher Degree	25.9%	21.1%	33.1%	B15002
Population Density (per sq mile)	1573.966	108.421	92.830*	B01003
% with Health Insurance Coverage	91.1%	92.7%	98.4%	B27001
% with Internet Access	80.5%	81.8%	85.8%	B28003

Table B.2: Split boundaries used to generate high-risk (treated) and low-risk (control) comparison groups based on ZIP code medians and U.S. national means of eight SDoH factors. Values that informed the choice of the split boundaries are bolded. For population density, we follow previous practices of urban-rural classification [327].

SDoH Factor	ZIP Code Median	U.S. Mean	Boundary	Treated	Control
Median Household Income	**$55,224**	$65,712	$55,224	≤ $55,224	> $55,224
% Unemployed	**3.0%**	2.9%	3.0%	≤ 3.0%	> 3.0%
% Black or African American Alone	1.6%	**12.8%**	12%	≥ 12%	< 12%
% Hispanic or Latino	4.0%	**18.4%**	18%	≥ 18%	< 18%
% with Bachelor or Higher Degree	**21.1%**	33.1%	21.1%	≤ 21.1%	> 21.1%
Population Density (per sq mile)	108.421	92.830	500	≥ 500	< 500
% with Health Insurance Coverage	**92.7%**	98.4%	92.7%	≤ 92.7%	> 92.7%
% with Internet Access	**81.8%**	85.8%	81.8%	≤ 81.8%	> 81.8%

Table B.3: Pearson correlation coefficients between SDoH factors across 25,801 ZIP codes in our dataset. These correlations reflect socioeconomic segregration and are controlled for through a matching-based approach (Methods). (* = attained Bachelor's degree or higher)

SDoH factor	% Unemp.	% Black	% Hispanic	% BA*	Pop. Density	% Insurance	% Internet
M. H. Income	-0.25	-0.23	-0.07	0.74	0.10	0.41	0.66
% Unemp.		0.32	0.15	-0.18	0.09	-0.23	-0.20
% Black			0.03	-0.09	0.15	-0.19	-0.23
% Hispanic				-0.07	0.21	-0.35	-0.04
% BA*					0.23	0.36	0.62
Pop. Density						0.00	0.10
% Insurance							0.41

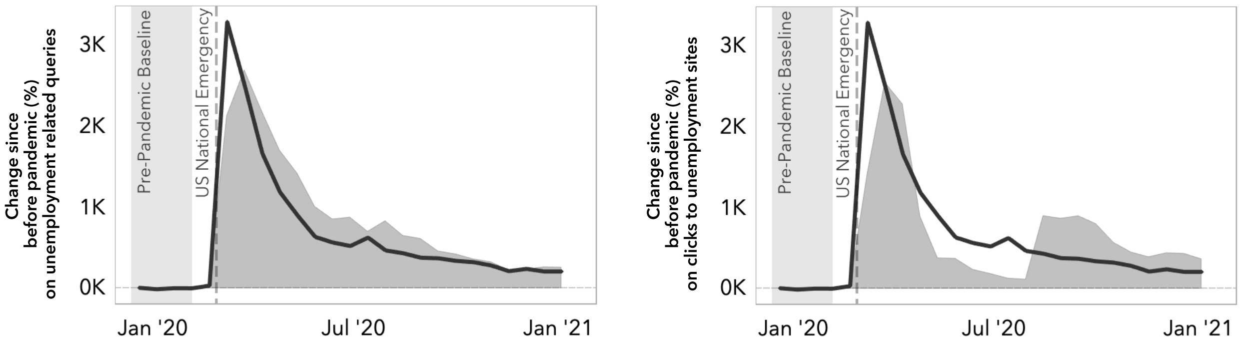
B.2 SEARCH CATEGORIES

Table B.4: Seven search categories examined and their example search strings and regular expressions. These categories were largely informed by our prior work [402]. In that prior work, we publicly shared the full list of our regular expressions in a github repository (https://github.com/jinasuh/pandemic_needs)

<table>
<tr><th>Search categories</th><th>Example search string</th><th>Example regular expression</th></tr>
<tr><td>Health condition related queries</td><td>arthritis; diabetes</td><td>\b(stroke|diabetes|covid|salmonella|autism|asthma|cancer)\b</td></tr>
<tr><td>Unemployment related queries</td><td>im unemployed; jobless benefits</td><td>\b((i m|im|i am) (unemployed|laid off|furloughed))\b</td></tr>
<tr><td>Clicks to state unemployment sites</td><td>click on www.michigan.gov/uia from search results</td><td>[\/.](www\.michigan\.gov\/uia|www\.in\.gov\/dwd\/2362\.htm)</td></tr>
<tr><td>Financial assistance related queries</td><td>relief fund; financial assistance</td><td>\b((relief|stimulus) (funds?|package|checks?))\b</td></tr>
<tr><td>Clicks to online learning sites</td><td>click on coursera.org from search results</td><td>[\/.](quizlet\.com|mooc-list\.com|coursera\.org|khanacademy\.org)</td></tr>
<tr><td>Food delivery related queries</td><td>grocery delivery; deliver meal</td><td>\b((food|grocery|meal) delivery)\b</td></tr>
<tr><td>Food assistance related queries</td><td>food stamps; snap program</td><td>b(food assistance|snap program|food stamps|food bank|food pantry)\b</td></tr>
</table>

B.3 UNEMPLOYMENT CLAIMS

Figure B.2: (a) Percent change in the unemployment related queries in Bing (shaded) and the reported unemployment claims from the US Department of Labor (line, https://oui.doleta.gov/unemploy/claims.asp) compared to pre-pandemic baseline. (b) Percent change in the clicks to unemployment sites in Bing (shaded) and the reported unemployment claims from the US Department of Labor (line, https://oui.doleta.gov/unemploy/claims.asp) compared to pre-pandemic baseline.



(a) Claims vs unemployment related queries

(b) Claims vs clicks to unemployment sites

B.4 MATCHING QUALITY

Table B.5: Summary of *pre-matching* balance assessment and sample sizes across SDoH factors. $|SMD| > 0.25$ indicates that the two ZIP code groups created along the split boundary are different and, thus, not comparable.

SDoH Factor	Split Boundary	Treated Count	Control Count	Max SMD	Mean SMD
Median Household Income	$55,224	12637	12637	1.018	0.459
% Unemployed	3.0%	12637	12637	0.385	0.269
% Black or Afr. Am. Alone	12%	4945	20329	0.467	0.300
% Hispanic or Latino	18%	4028	21246	0.591	0.218
% with Bachelor or Higher	21.1%	12637	12637	0.975	0.416
Population Density (per mi^2)	500	8037	17237	0.689	0.392
% with Health Ins. Cov.	92.7%	12666	12608	0.629	0.396
% with Internet Access	81.8%	12637	12637	0.974	0.429

Table B.6: Summary of *post-matching* balance assessment and sample sizes across SDoH factors for treated and control groups.

SDoH Factor	Max SMD	Mean SMD	Tr. Matched	Tr. Unmat.	Ct. Matched	Ct. Unmat.
Median Household Income	0.236	0.105	12555	21	3854	8720
% Unemployed	0.088	0.026	12544	31	5917	6658
% Black or Afr. Am. Alone	0.144	0.072	4932	0	1603	18615
% Hispanic or Latino	0.131	0.082	4018	0	2247	18885
% with Bachelor or Higher	0.113	0.069	12572	3	3981	8594
Population Density (per mi^2)	0.181	0.114	8022	6	2836	14286
% with Health Ins. Cov.	0.067	0.030	8508	0	4339	12303
% with Internet Access	0.176	0.075	12575	0	3623	8952

Table B.7: Balance assessment between unmatched high (Treated) and low (Control) '% Black or African American Alone' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.590	0.100	0.490	1.606
Median Household Income	49810.138	63497.945	-13687.808	-0.680
Unemployed	0.045	0.031	0.015	0.628
Hispanic or Latino	0.125	0.095	0.030	0.201
With Bachelor or Higher Degree Attained	0.232	0.265	-0.034	-0.244
Population Density (per sq mile)	2975.905	1172.126	1803.778	0.245
With Health Insurance Coverage	0.883	0.918	-0.034	-0.575
With Internet Access	0.766	0.816	-0.050	-0.414
Black or African American Alone	0.337	0.021	0.316	1.517

Table B.8: Balance assessment between matched high (Treated) and low (Control) '% Black or African American Alone' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.887	0.791	0.096	0.043
Median Household Income	49835.515	52509.535	-2674.020	-0.133
Unemployed	0.045	0.043	0.002	0.097
Hispanic or Latino	0.125	0.126	-0.001	-0.004
With Bachelor or Higher Degree Attained	0.232	0.240	-0.008	-0.059
Population Density (per sq mile)	2981.898	2250.314	731.584	0.099
With Health Insurance Coverage	0.884	0.881	0.003	0.044
With Internet Access	0.766	0.775	-0.009	-0.072
Black or African American Alone	0.337	0.041	0.297	1.425

Table B.9: Balance assessment between unmatched high (Treated) and low (Control) '% Hispanic or Latino' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.368	0.120	0.249	1.023
Median Household Income	57217.731	61502.769	-4285.038	-0.209
Unemployed	0.041	0.032	0.009	0.398
Black or African American Alone	0.099	0.080	0.019	0.145
With Bachelor or Higher Degree Attained	0.234	0.264	-0.030	-0.224
Population Density (per sq mile)	4021.757	1051.698	2970.060	0.333
With Health Insurance Coverage	0.864	0.920	-0.056	-0.751
With Internet Access	0.806	0.806	-0.000	-0.001
Hispanic or Latino	0.399	0.044	0.355	1.736

Table B.10: Balance assessment between matched high (Treated) and low (Control) '% Hispanic or Latino' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	-0.639	-0.699	0.060	0.042
Median Household Income	57234.061	58095.893	-861.832	-0.042
Unemployed	0.041	0.038	0.002	0.105
Black or African American Alone	0.099	0.085	0.014	0.111
With Bachelor or Higher Degree Attained	0.234	0.239	-0.006	-0.042
Population Density (per sq mile)	4029.507	2986.171	1043.336	0.117
With Health Insurance Coverage	0.864	0.865	-0.001	-0.018
With Internet Access	0.806	0.817	-0.010	-0.096
Hispanic or Latino	0.399	0.086	0.313	1.530

Table B.11: Balance assessment between unmatched high (Control) and low (Treated) 'Median Household Income' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.725	0.275	0.450	1.855
Unemployed	0.039	0.028	0.011	0.412
Black or African American Alone	0.117	0.049	0.068	0.343
Hispanic or Latino	0.110	0.091	0.019	0.102
With Bachelor or Higher Degree Attained	0.177	0.341	-0.164	-1.751
Population Density (per sq mile)	1266.027	1784.065	-518.038	-0.107
With Health Insurance Coverage	0.887	0.934	-0.047	-0.692
With Internet Access	0.744	0.868	-0.125	-1.257
Median Household Income	43356.749	78282.949	-34926.200	-4.189

Table B.12: Balance assessment between matched high (Control) and low (Treated) 'Median Household Income' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	1.567	1.449	0.119	0.060
Unemployed	0.039	0.036	0.003	0.113
Black or African American Alone	0.117	0.092	0.025	0.126
Hispanic or Latino	0.111	0.091	0.020	0.107
With Bachelor or Higher Degree Attained	0.177	0.180	-0.003	-0.029
Population Density (per sq mile)	1272.878	747.235	525.643	0.108
With Health Insurance Coverage	0.888	0.889	-0.001	-0.021
With Internet Access	0.745	0.746	-0.001	-0.011
Median Household Income	43401.084	62615.121	-19214.037	-2.306

Table B.13: Balance assessment between unmatched high (Control) and low (Treated) '% with Bachelor or Higher Degree Attained' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.714	0.286	0.429	1.869
Median Household Income	47879.339	73760.359	-25881.020	-2.046
Unemployed	0.037	0.029	0.008	0.308
Black or African American Alone	0.095	0.072	0.023	0.125
Hispanic or Latino	0.109	0.093	0.016	0.086
Population Density (per sq mile)	788.145	2261.947	-1473.803	-0.441
With Health Insurance Coverage	0.890	0.931	-0.041	-0.571
With Internet Access	0.745	0.867	-0.121	-1.222
With Bachelor or Higher Degree Attained	0.141	0.376	-0.235	-5.334

Table B.14: Balance assessment between matched high (Control) and low (Treated) '% with Bachelor or Higher Degree Attained' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	1.336	1.242	0.094	0.058
Median Household Income	47895.952	48454.763	-558.811	-0.044
Unemployed	0.037	0.036	0.002	0.073
Black or African American Alone	0.095	0.084	0.011	0.060
Hispanic or Latino	0.109	0.097	0.013	0.066
Population Density (per sq mile)	793.735	751.143	42.592	0.013
With Health Insurance Coverage	0.891	0.897	-0.007	-0.097
With Internet Access	0.746	0.751	-0.005	-0.050
With Bachelor or Higher Degree Attained	0.142	0.269	-0.128	-2.903

Table B.15: Balance assessment between unmatched high (Treated) and low (Control) 'Population Density (per sq mile)' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.645	0.166	0.479	1.701
Median Household Income	71129.001	56013.058	15115.943	0.462
Unemployed	0.038	0.031	0.006	0.322
Black or African American Alone	0.144	0.055	0.089	0.450
Hispanic or Latino	0.169	0.069	0.100	0.522
With Bachelor or Higher Degree Attained	0.364	0.210	0.154	0.805
With Health Insurance Coverage	0.915	0.909	0.006	0.097
With Internet Access	0.860	0.781	0.079	0.877
Population Density (per sq mile)	4599.740	91.425	4508.315	0.510

Table B.16: Balance assessment between matched high (Treated) and low (Control) ‘Population Density (per sq mile)’ groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.882	0.763	0.118	0.067
Median Household Income	71169.714	72899.376	-1729.662	-0.053
Unemployed	0.038	0.035	0.002	0.132
Black or African American Alone	0.144	0.124	0.020	0.099
Hispanic or Latino	0.169	0.154	0.015	0.079
With Bachelor or Higher Degree Attained	0.363	0.371	-0.007	-0.038
With Health Insurance Coverage	0.915	0.925	-0.010	-0.156
With Internet Access	0.860	0.873	-0.013	-0.147
Population Density (per sq mile)	4598.958	227.548	4371.410	0.494

Table B.17: Balance assessment between unmatched high (Treated) and low (Control) ‘% Unemployed’ groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.581	0.419	0.162	0.781
Median Household Income	55178.441	66461.257	-11282.816	-0.503
Black or African American Alone	0.125	0.042	0.083	0.421
Hispanic or Latino	0.132	0.070	0.062	0.328
With Bachelor or Higher Degree Attained	0.230	0.288	-0.058	-0.416
Population Density (per sq mile)	2062.020	988.072	1073.948	0.173
With Health Insurance Coverage	0.898	0.924	-0.026	-0.404
With Internet Access	0.791	0.821	-0.029	-0.272
Unemployed	0.049	0.018	0.030	1.401

Table B.18: Balance assessment between matched high (Treated) and low (Control) ‘% Unemployed’ groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.460	0.423	0.037	0.032
Median Household Income	55192.003	55360.831	-168.828	-0.008
Black or African American Alone	0.124	0.122	0.002	0.010
Hispanic or Latino	0.131	0.125	0.006	0.030
With Bachelor or Higher Degree Attained	0.230	0.225	0.004	0.030
Population Density (per sq mile)	1902.049	1476.827	425.222	0.069
With Health Insurance Coverage	0.898	0.898	-0.001	-0.010
With Internet Access	0.791	0.791	0.001	0.006
Unemployed	0.048	0.021	0.028	1.303

Table B.19: Balance assessment between unmatched high (Control) and low (Treated) '% with Health Insurance Coverage' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.528	0.242	0.286	1.140
Median Household Income	48183.387	67291.556	-19108.168	-1.336
Unemployed	0.040	0.030	0.010	0.374
Black or African American Alone	0.133	0.058	0.075	0.378
Hispanic or Latino	0.167	0.067	0.101	0.458
With Bachelor or Higher Degree Attained	0.187	0.296	-0.109	-1.076
Population Density (per sq mile)	1529.525	1522.752	6.774	0.001
With Internet Access	0.752	0.834	-0.082	-0.737
With Health Insurance Coverage	0.841	0.947	-0.106	-1.770

Table B.20: Balance assessment between matched high (Control) and low (Treated) '% with Health Insurance Coverage' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.255	0.199	0.055	0.038
Median Household Income	48177.005	48170.216	6.789	0.000
Unemployed	0.040	0.039	0.001	0.025
Black or African American Alone	0.133	0.126	0.007	0.035
Hispanic or Latino	0.168	0.158	0.009	0.043
With Bachelor or Higher Degree Attained	0.187	0.187	0.000	0.005
Population Density (per sq mile)	1537.624	1250.643	286.982	0.060
With Internet Access	0.752	0.753	-0.001	-0.008
With Health Insurance Coverage	0.841	0.932	-0.092	-1.530

Table B.21: Balance assessment between unmatched high (Control) and low (Treated) '% with Internet Access' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	0.749	0.251	0.497	2.260
Median Household Income	46716.098	74923.599	-28207.501	-2.274
Unemployed	0.037	0.030	0.007	0.256
Black or African American Alone	0.107	0.060	0.047	0.244
Hispanic or Latino	0.102	0.099	0.003	0.019
With Bachelor or Higher Degree Attained	0.168	0.350	-0.182	-2.326
Population Density (per sq mile)	1003.589	2046.502	-1042.913	-0.225
With Health Insurance Coverage	0.889	0.932	-0.043	-0.603
With Internet Access	0.723	0.889	-0.166	-1.980

Table B.22: Balance assessment between matched high (Control) and low (Treated) '% with Internet Access' groups

SDoH Factor	Means Treated	Means Control	Mean Diff.	Std. Mean Diff.
distance	1.436	1.234	0.202	0.136
Median Household Income	46721.584	48710.356	-1988.773	-0.160
Unemployed	0.037	0.035	0.001	0.056
Black or African American Alone	0.107	0.096	0.011	0.059
Hispanic or Latino	0.103	0.093	0.009	0.050
With Bachelor or Higher Degree Attained	0.168	0.173	-0.005	-0.066
Population Density (per sq mile)	1009.178	898.001	111.178	0.024
With Health Insurance Coverage	0.889	0.893	-0.003	-0.049
With Internet Access	0.724	0.854	-0.130	-1.559

B.5 SEARCH TRENDS

Figure B.3: Normalized query volumes across five categories in years 2019 and 2020 across Bing and Google.

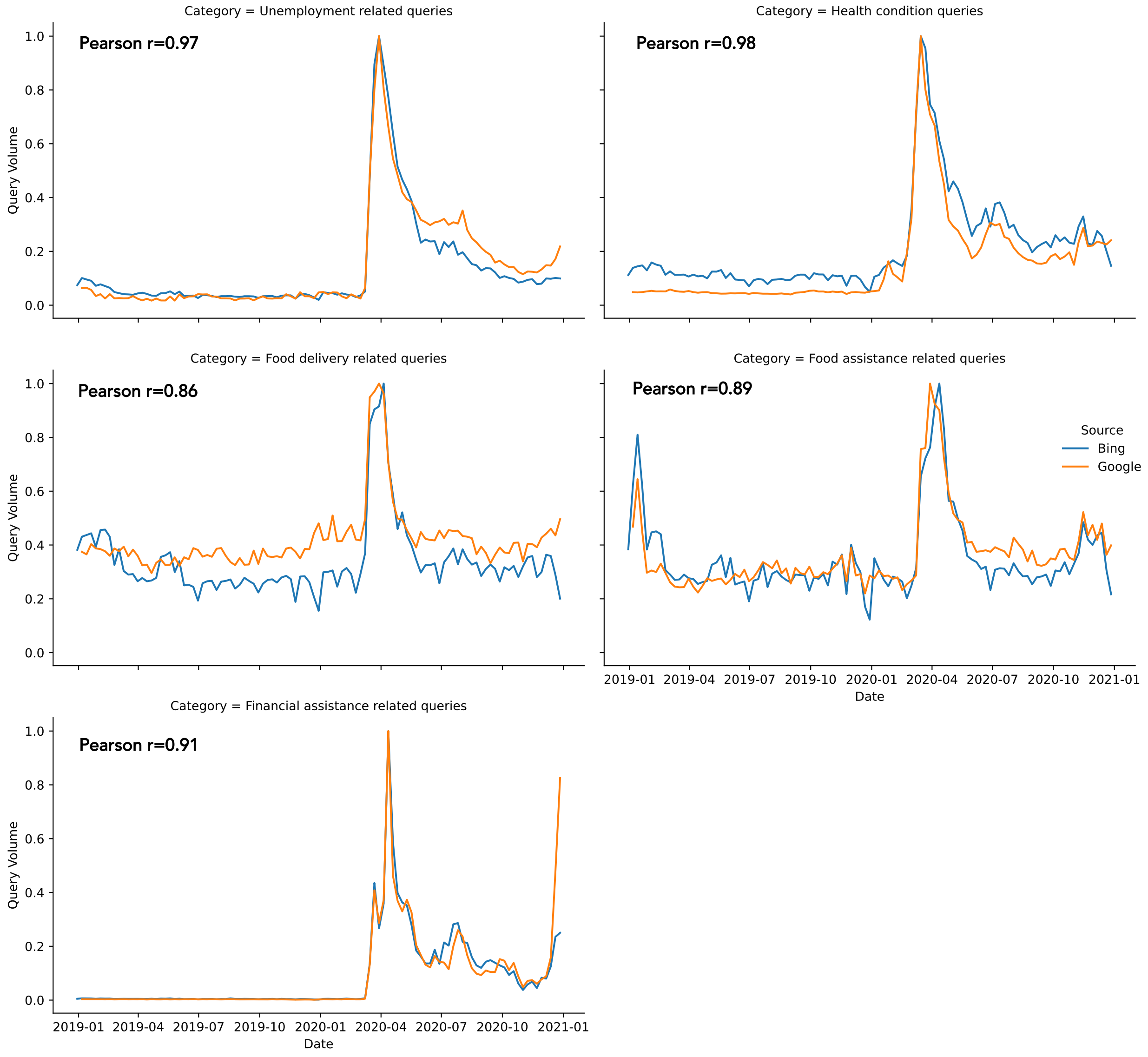


Figure B.4: Frequency of search keywords “covid”, “coronavirus”, “mers”, “gonorrhea”, and “zika” in 2019, scaled to the maximum frequency within the data presented in the figure.

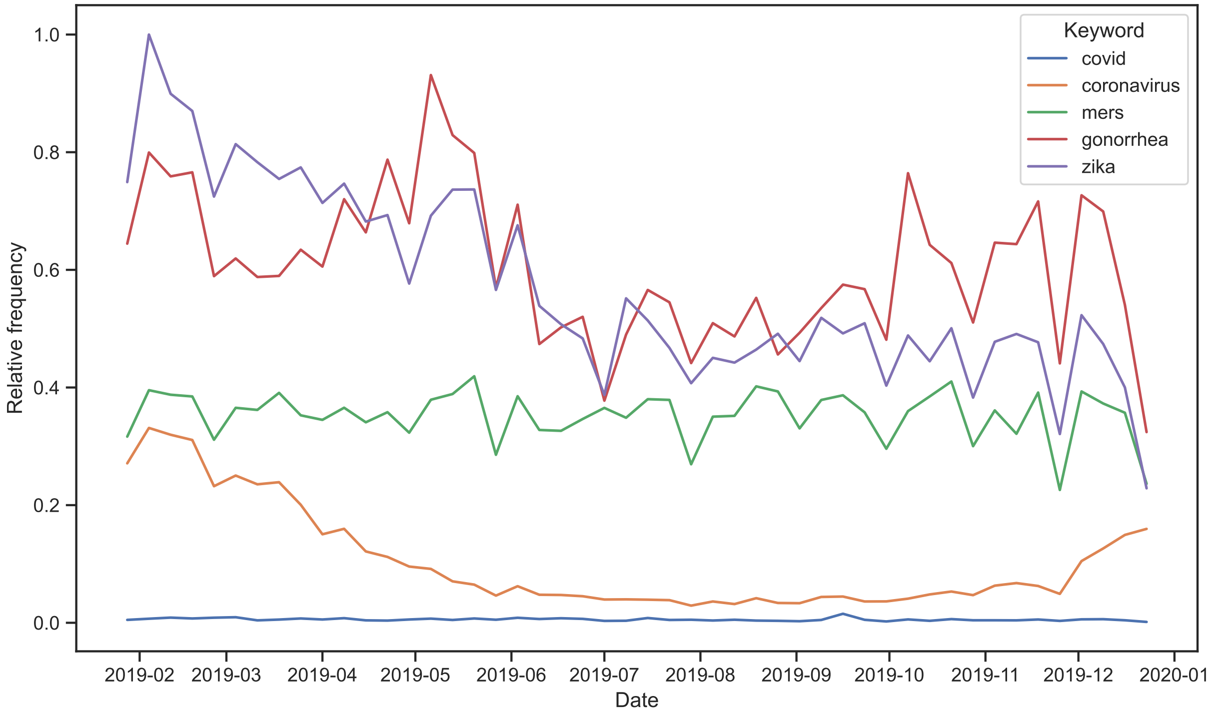
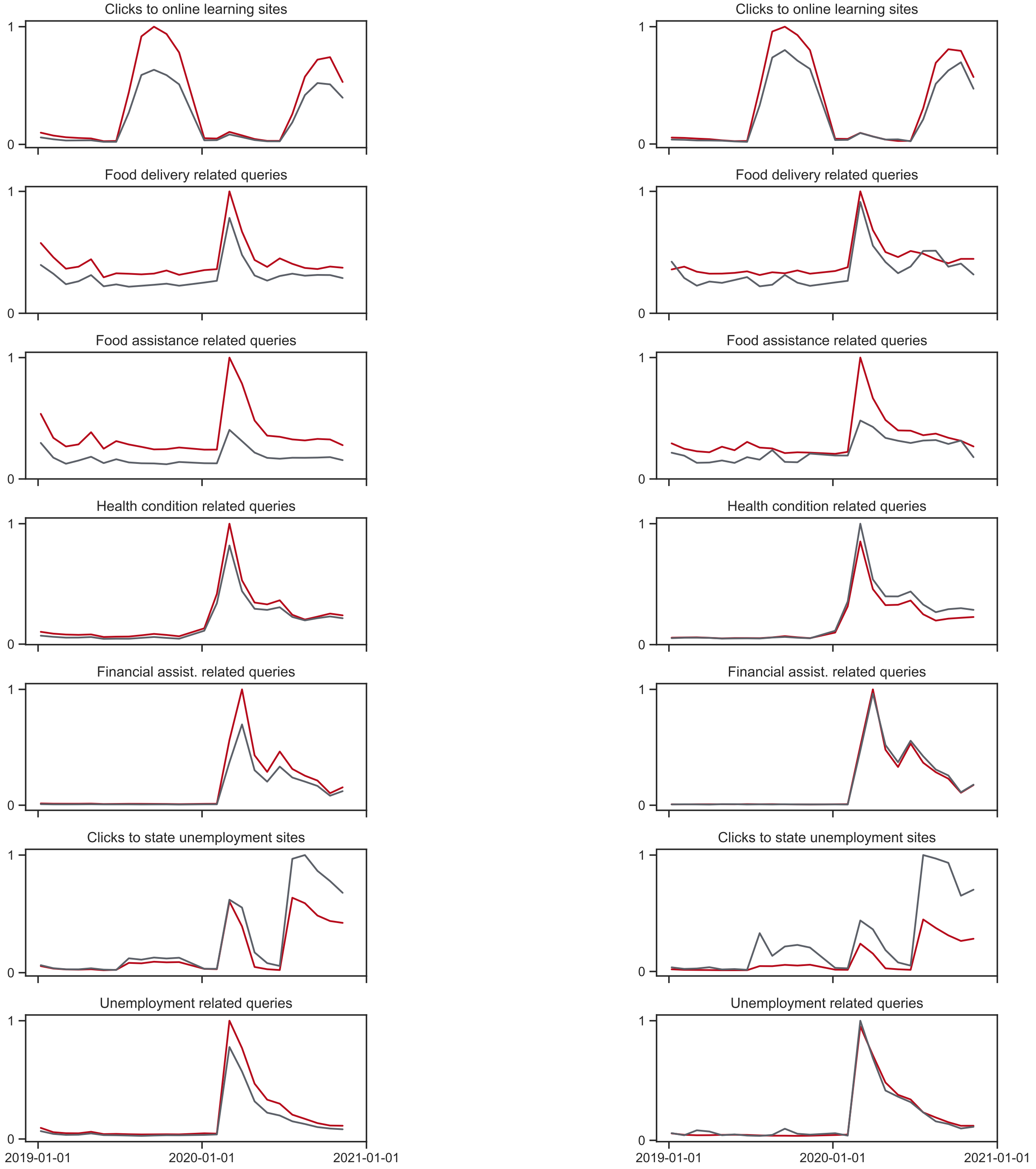


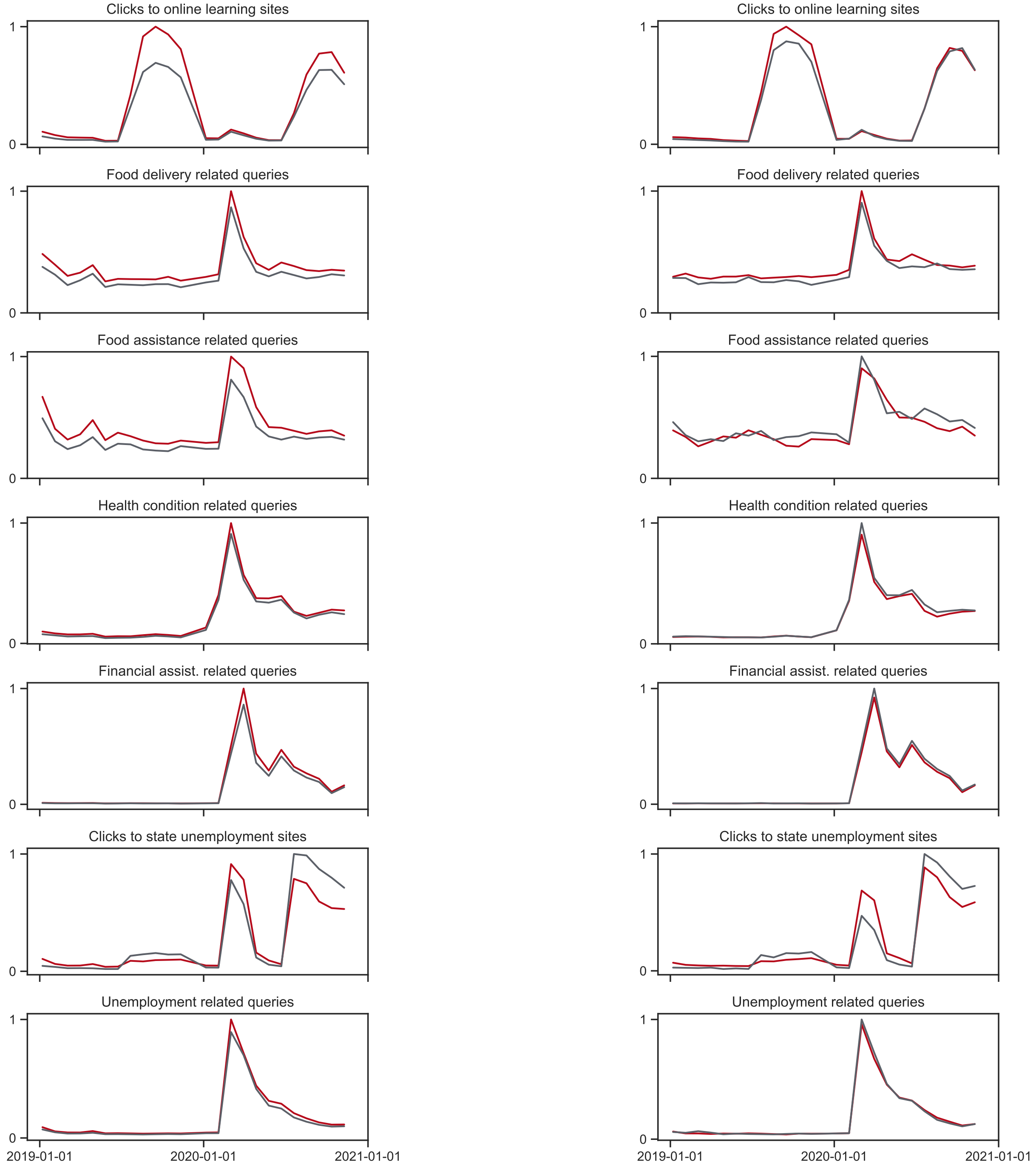
Figure B.5: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with % of the population with Black or African American alone $\geq 12\%$, and gray lines indicate ZIP codes with % of the population with Black or African American alone $< 12\%$.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

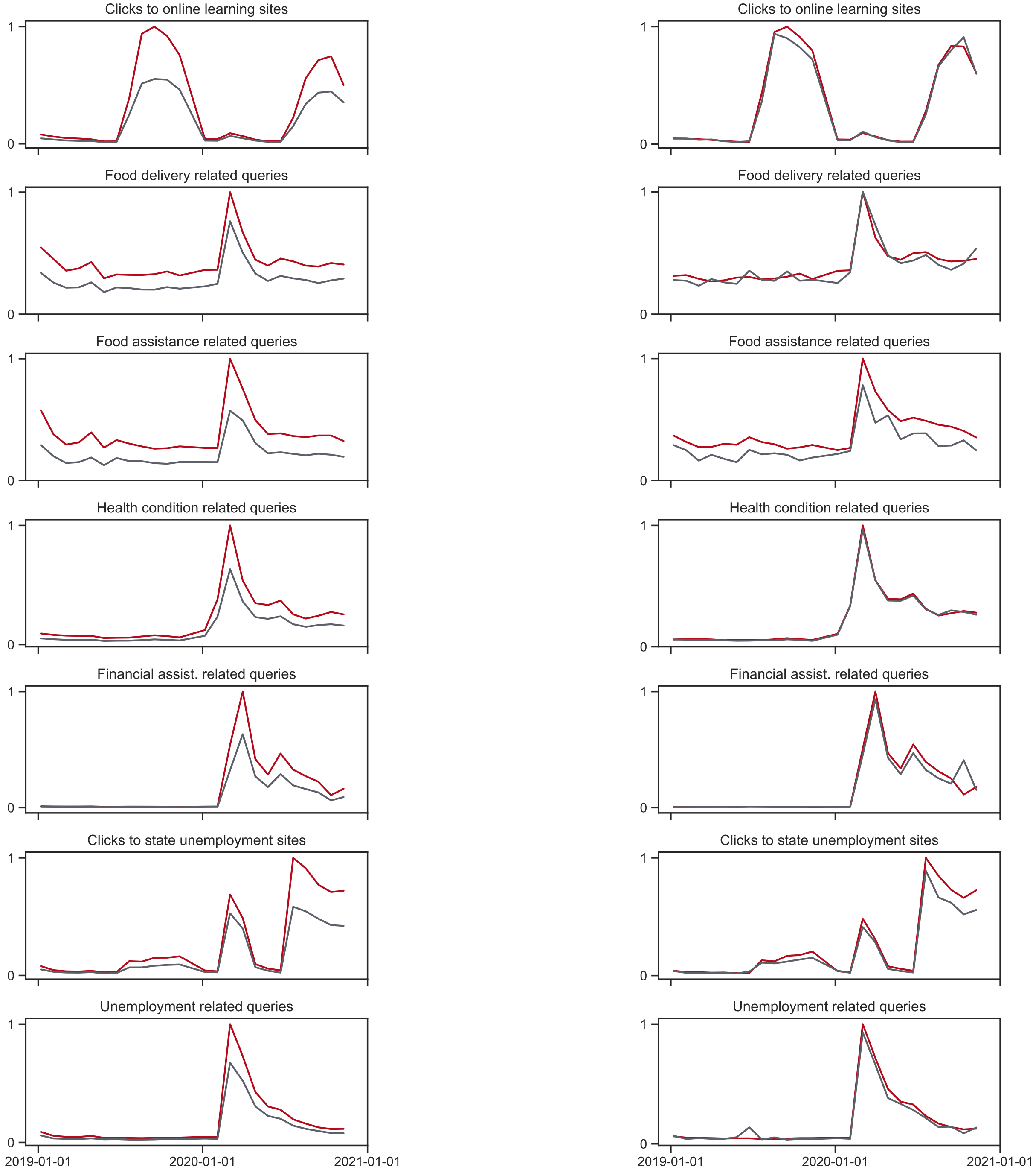
Figure B.6: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with % of the population with Hispanic origin $\geq 18\%$, and gray lines indicate ZIP codes with % of the population with Hispanic origin $< 18\%$.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

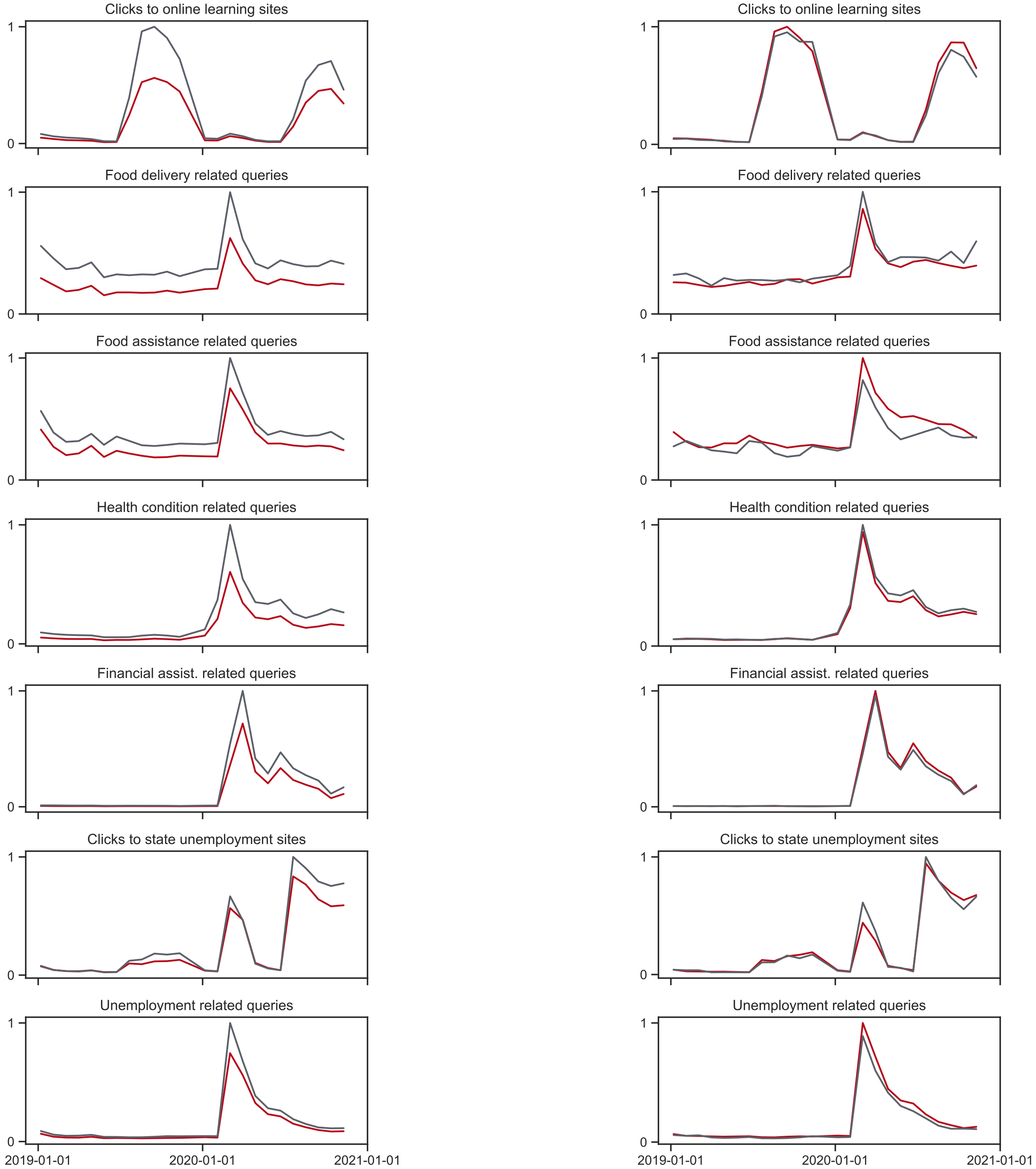
Figure B.7: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with median household income ≤ \$55,224, and gray lines indicate ZIP codes with median household income > \$55,224.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

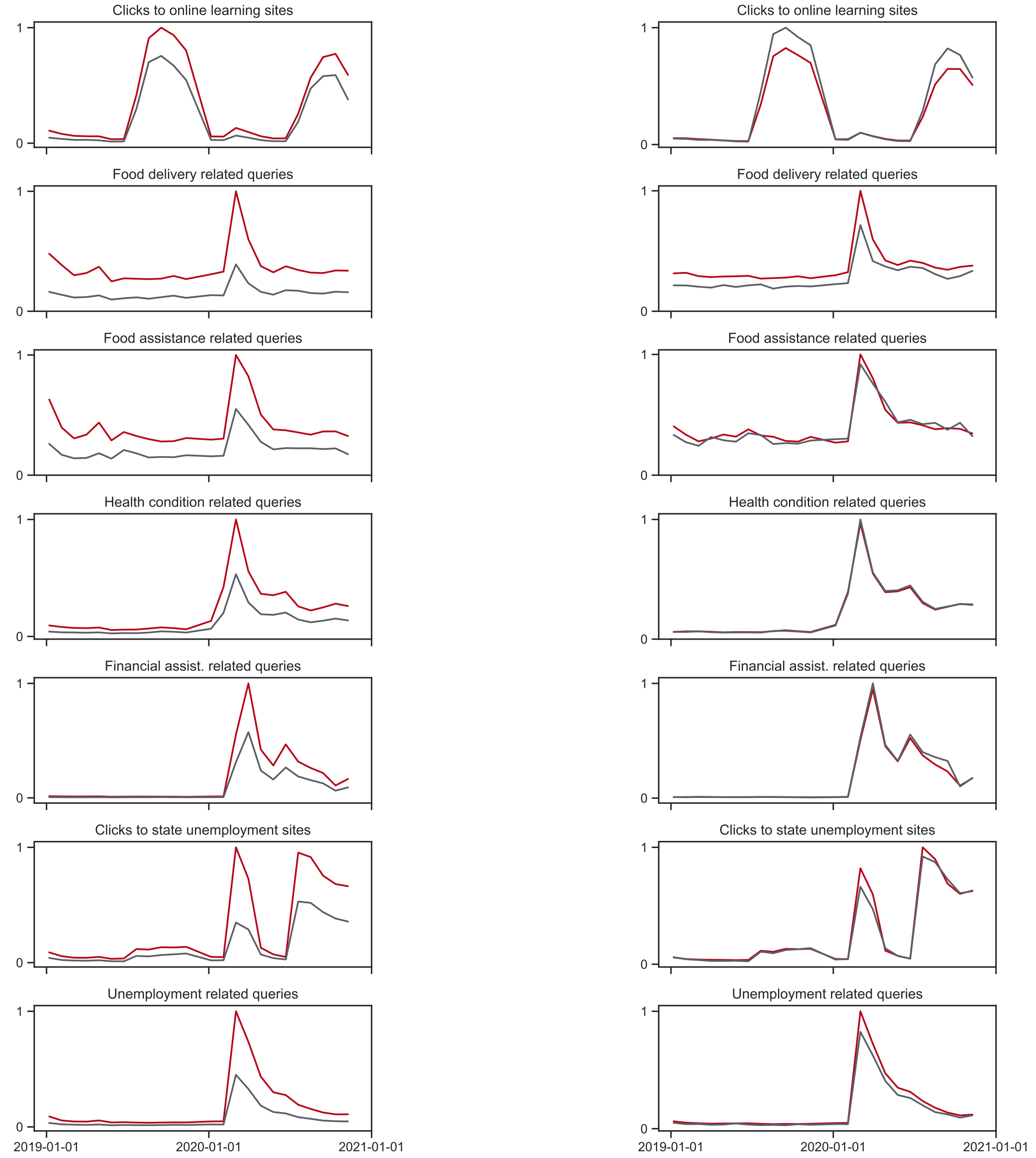
Figure B.8: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with % of the population with Bachelor's degree or higher degrees $\leq 21\%$, and gray lines indicate ZIP codes with Bachelor's degree or higher degrees $> 21\%$.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

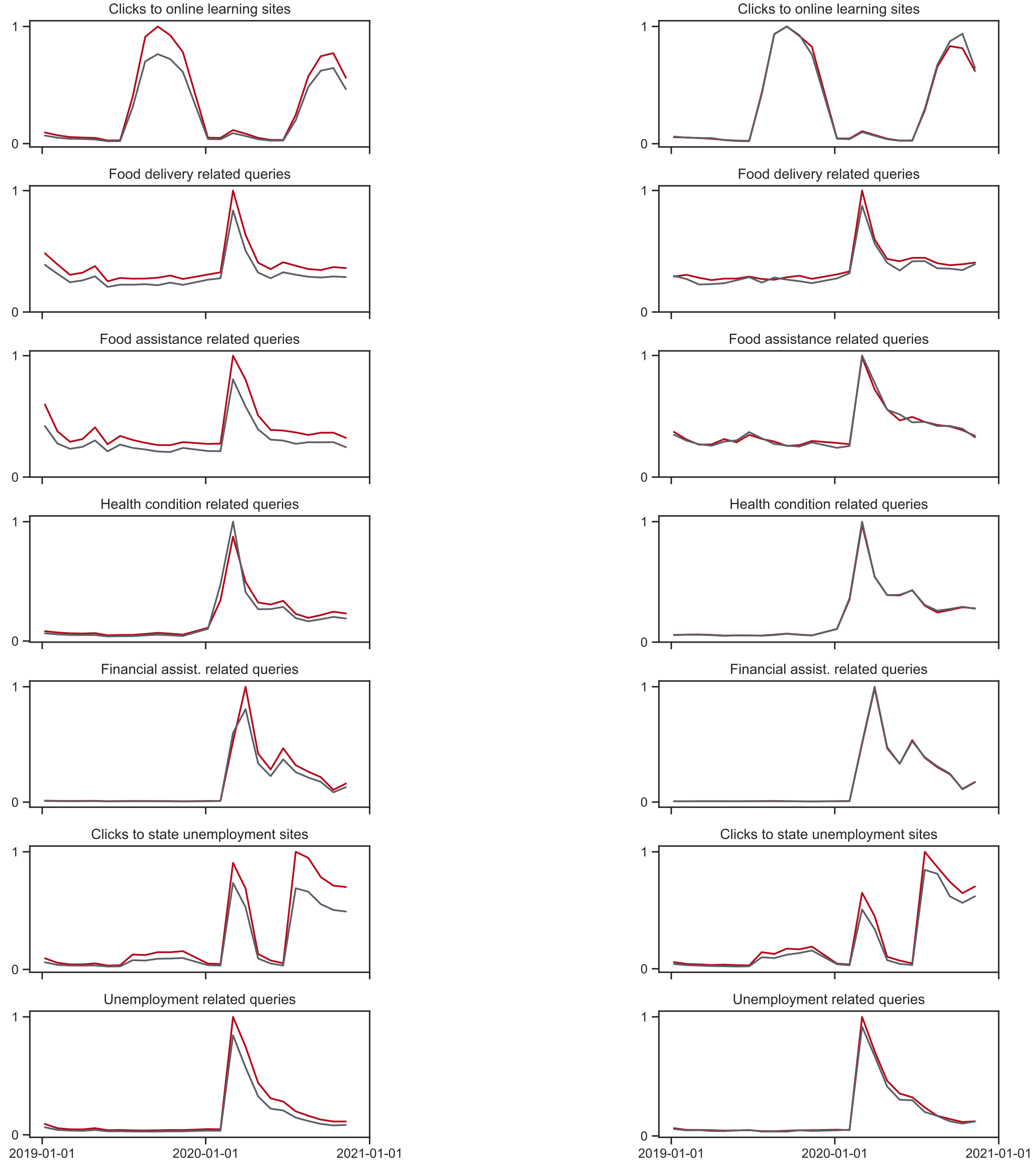
Figure B.9: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with population density $\geq$ 500 people per square mile, and gray lines indicate ZIP codes with internet access $<$ 500 people per square mile.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

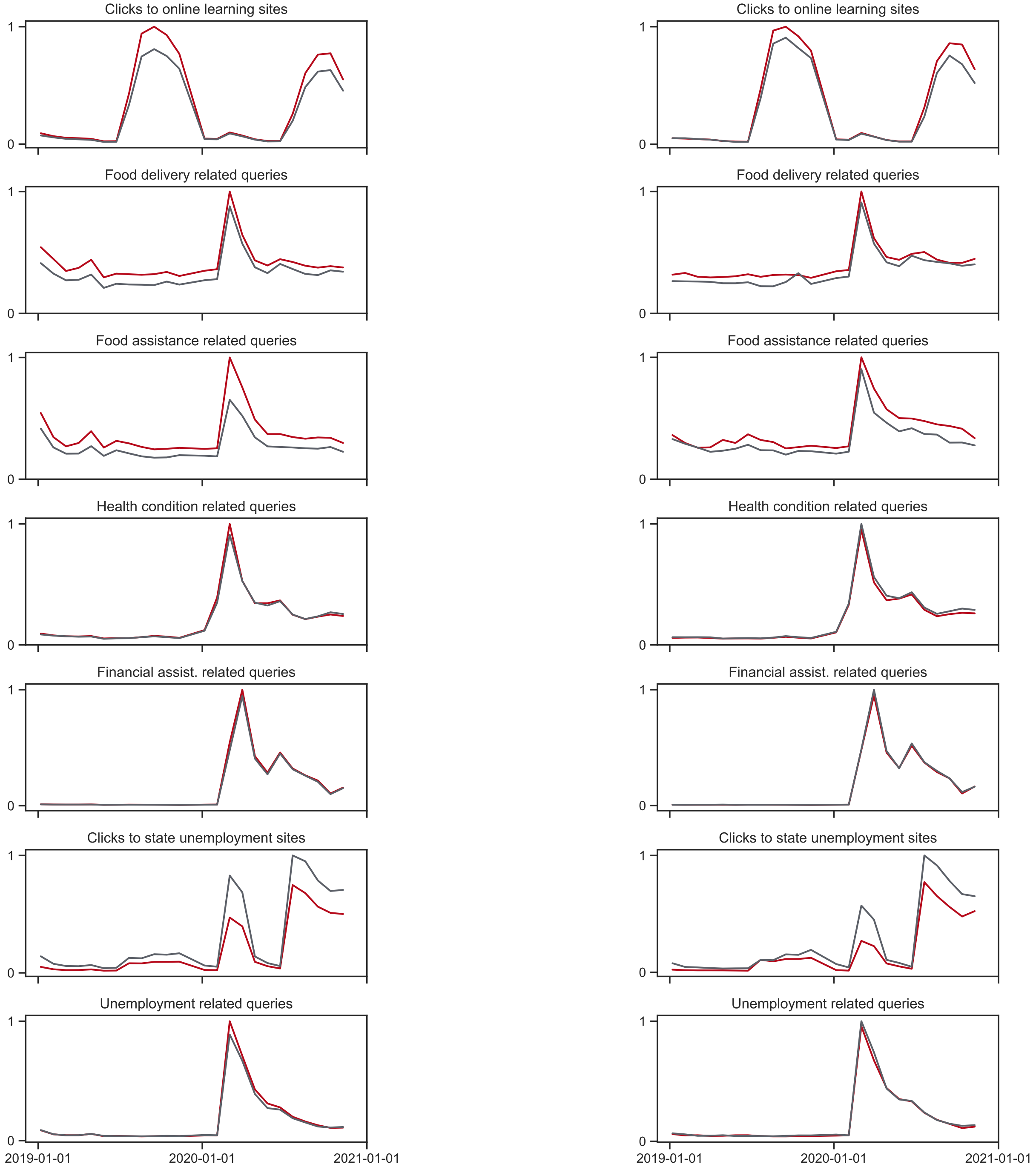
Figure B.10: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with % of the population unemployed $\geq 3\%$, and gray lines indicate ZIP codes with % of the population unemployed $< 3\%$.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

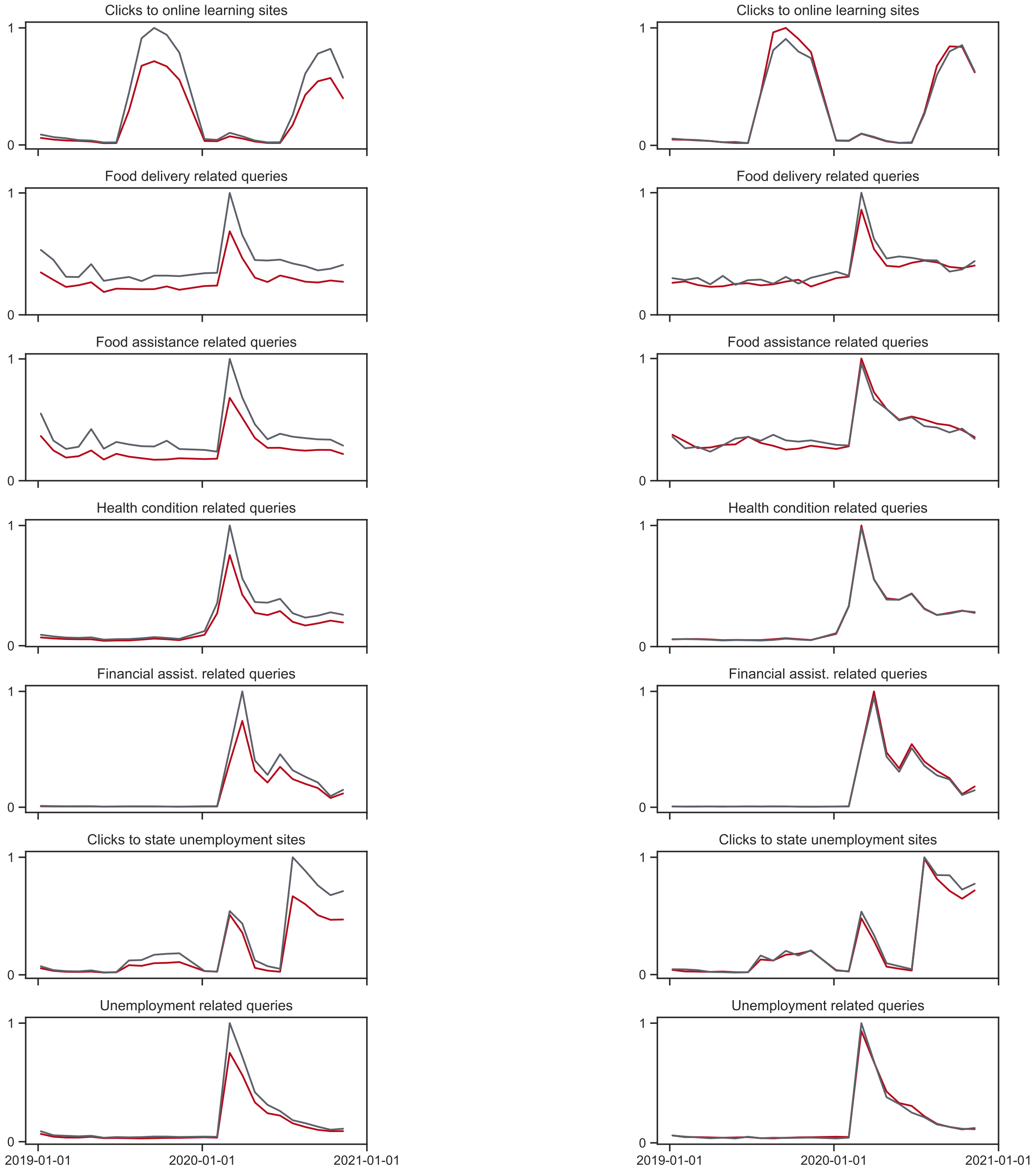
Figure B.11: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with % of the population with health insurance coverage ≤ 93%, and gray lines indicate ZIP codes with health insurance coverage > 93%.



(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

Figure B.12: (a) Average query frequencies per ZIP code across seven search categories, and (b) average proportion of total queries per ZIP code across seven search categories. Query frequencies and proportions are scaled to the maximum of 1 within each figure. Red lines indicate ZIP codes with % of the population with internet access $\leq 82\%$, and gray lines indicate ZIP codes with internet access $> 82\%$.



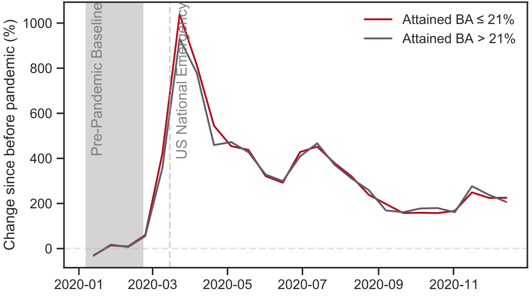
(a) Average query frequencies per ZIP code

(b) Average query proportions per ZIP code

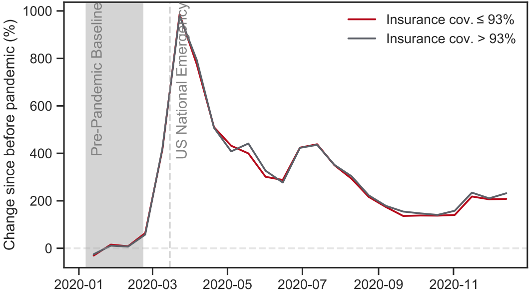
B.6 RELATIVE DIFFERENCES ACROSS GROUPS

Figure B.13: Percent change in ‘Health condition related queries’ between two matched groups across eight SDoH factors.

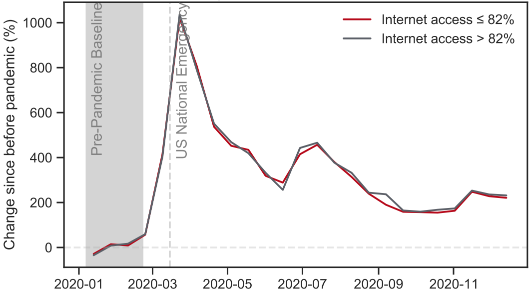
(a) Percent change in 'Health condition related queries' between two matched groups across 'Education level'



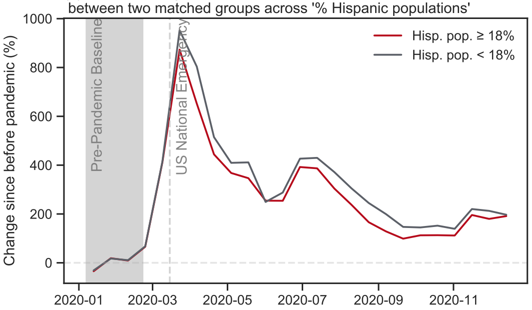
(b) Percent change in 'Health condition related queries' between two matched groups across '% Insurance coverage'



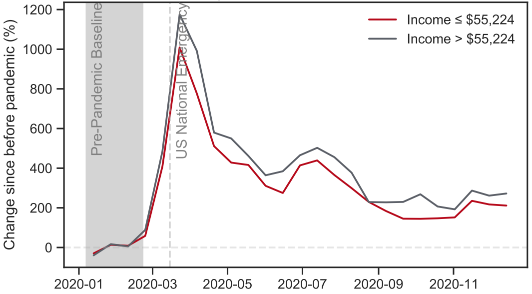
(c) Percent change in 'Health condition related queries' between two matched groups across '% Internet access'



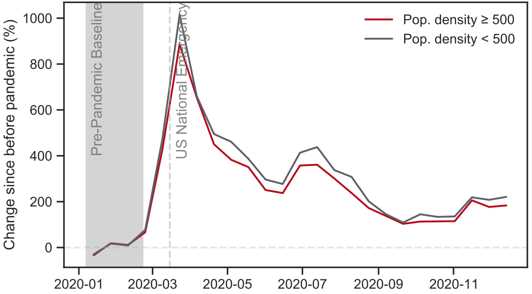
(d) Percent change in 'Health condition related queries' between two matched groups across '% Hispanic populations'



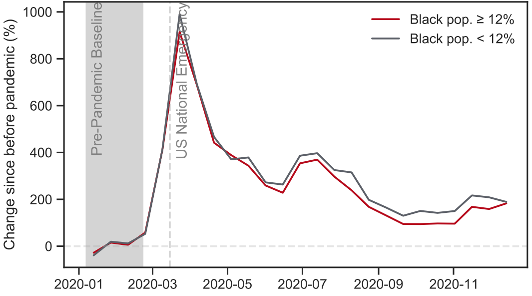
(e) Percent change in 'Health condition related queries' between two matched groups across 'Income'



(f) Percent change in 'Health condition related queries' between two matched groups across 'Population density'



(g) Percent change in 'Health condition related queries' between two matched groups across '% Black population'



(h) Percent change in 'Health condition related queries' between two matched groups across '% Unemployed'

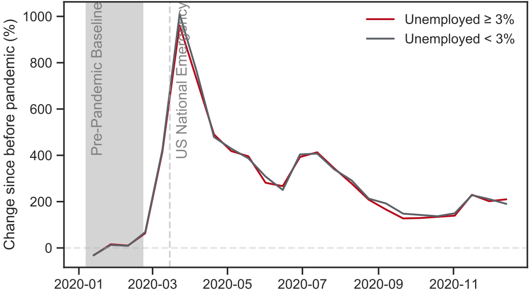
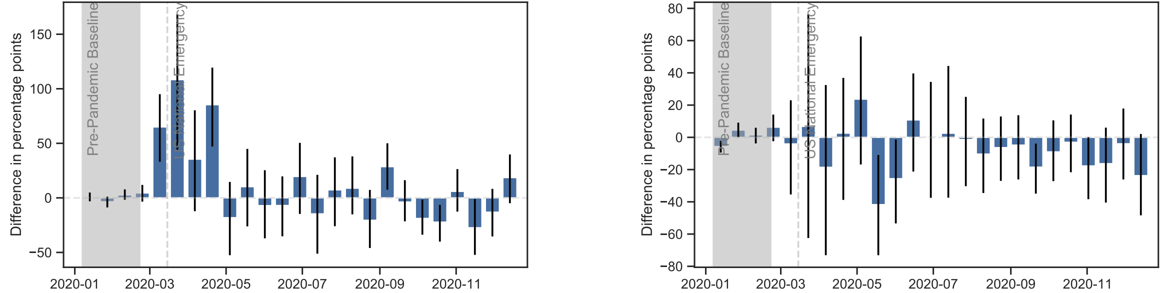
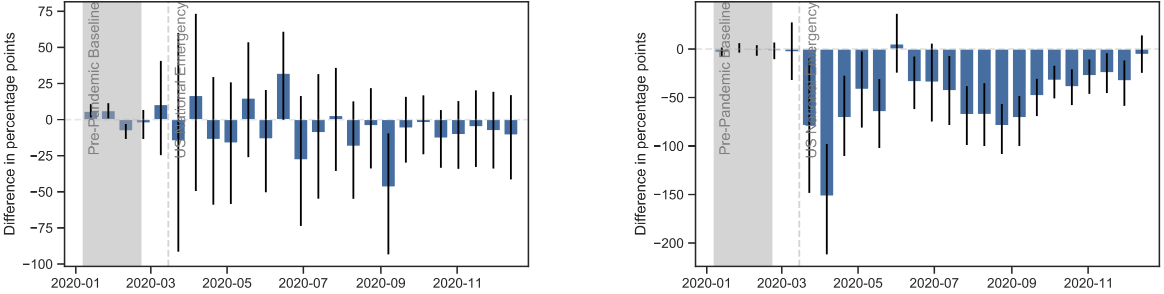


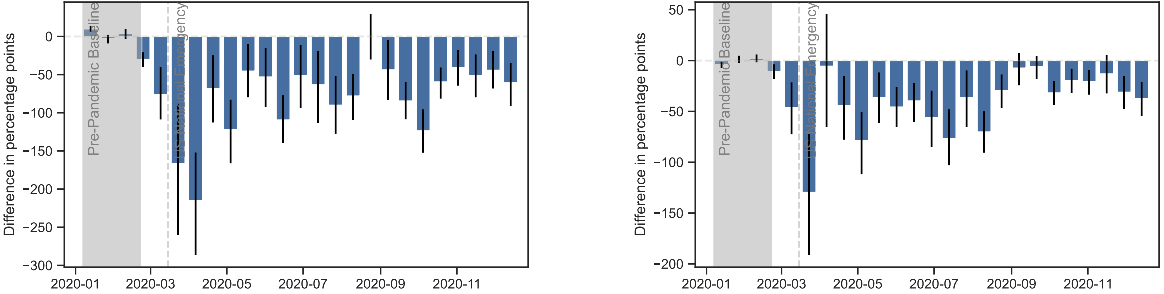
Figure B.14: Differences in percentage points for changes in 'Health condition related queries' between two matched groups across eight SDoH factors.


(a) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across 'Education level'.
High risk (Attained BA ≤ 21%) - Low risk (Attained BA > 21%)

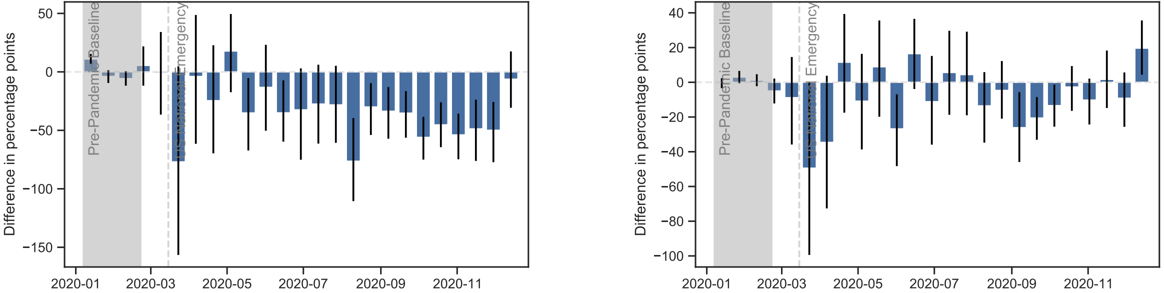
(b) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across '% Insurance coverage'.
High risk (Insurance cov. ≤ 93%) - Low risk (Insurance cov. > 93%)


(c) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across '% Internet access'.
High risk (Internet access ≤ 82%) - Low risk (Internet access > 82%)

(d) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across '% Hispanic populations'.
High risk (Hisp. pop. ≥ 18%) - Low risk (Hisp. pop. < 18%)

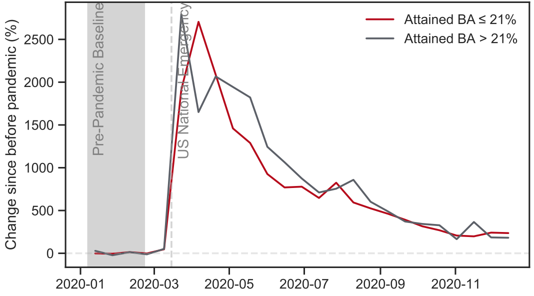

(e) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across 'Income'.
High risk (Income ≤ $55,224) - Low risk (Income > $55,224)

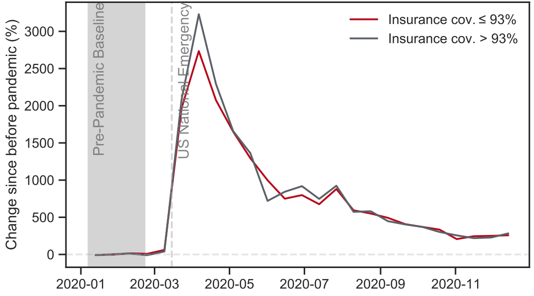
(f) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across 'Population density'.
High risk (Pop. density ≥ 500) - Low risk (Pop. density < 500)

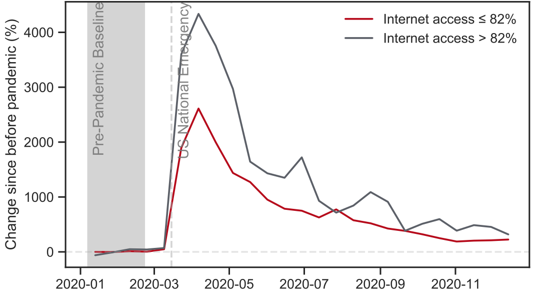

(g) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across '% Black population'.
High risk (Black pop. ≥ 12%) - Low risk (Black pop. < 12%)

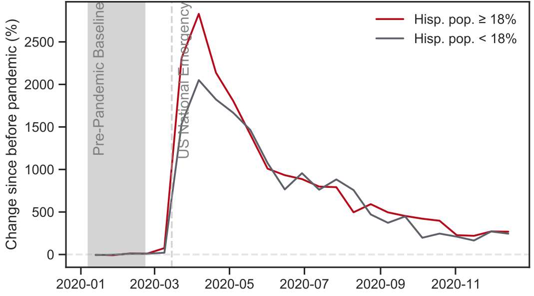
(h) Diff. in percentage points for changes in 'Health condition related queries' between two matched groups across '% Unemployed'.
High risk (Unemployed ≥ 3%) - Low risk (Unemployed < 3%)

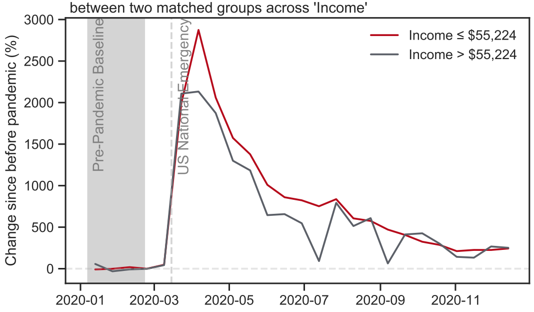
Figure B.15: Percent change in ‘Unemployment related queries’ between two matched groups across eight SDoH factors.

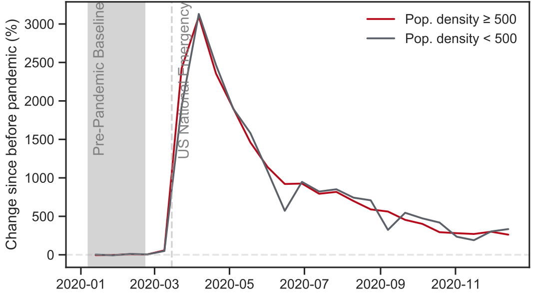

(a) Percent change in 'Unemployment related queries' between two matched groups across 'Education level'

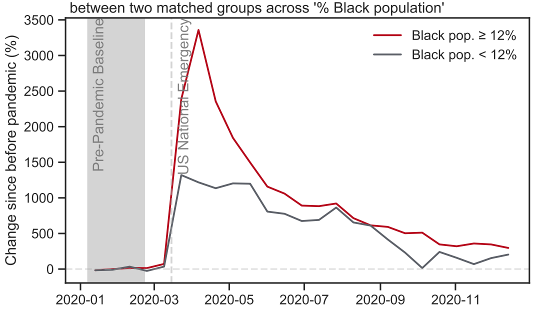

(b) Percent change in 'Unemployment related queries' between two matched groups across '% Insurance coverage'


(c) Percent change in 'Unemployment related queries' between two matched groups across '% Internet access'


(d) Percent change in 'Unemployment related queries' between two matched groups across '% Hispanic populations'


(e) Percent change in 'Unemployment related queries' between two matched groups across 'Income'


(f) Percent change in 'Unemployment related queries' between two matched groups across 'Population density'


(g) Percent change in 'Unemployment related queries' between two matched groups across '% Black population'

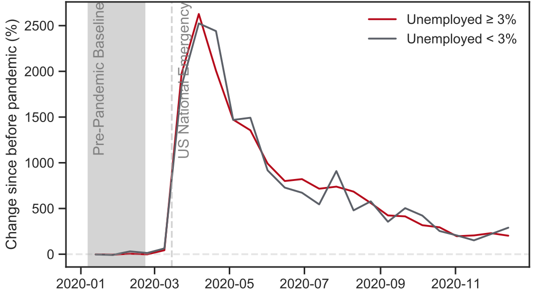
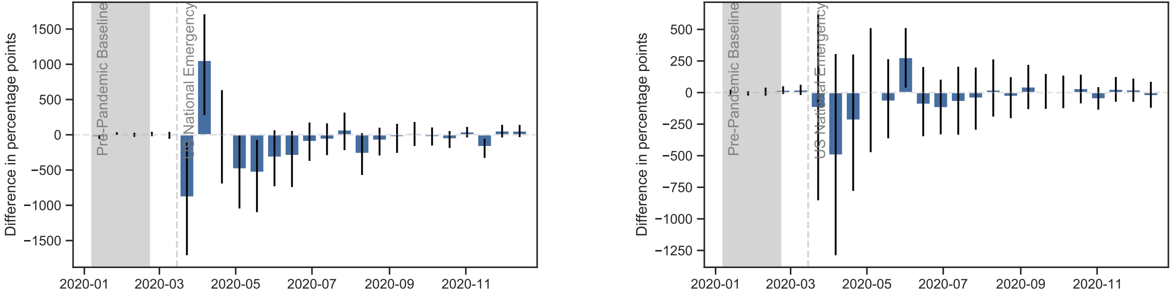
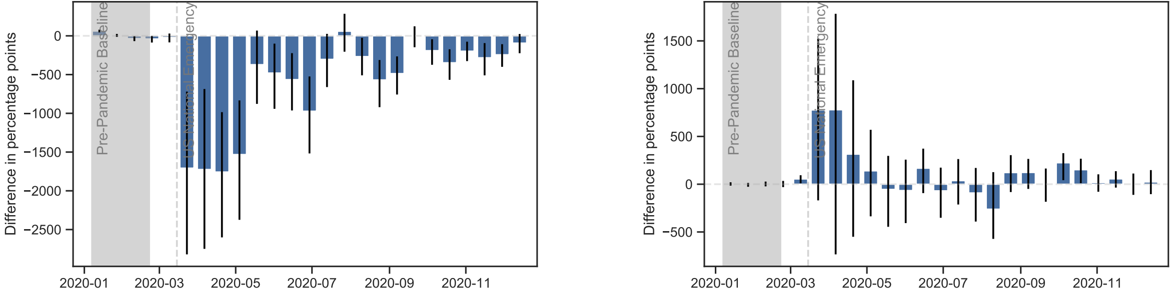

(h) Percent change in 'Unemployment related queries' between two matched groups across '% Unemployed'

Figure B.16: Differences in percentage points for changes in ‘Unemployment related queries’ between two matched groups across eight SDoH factors.



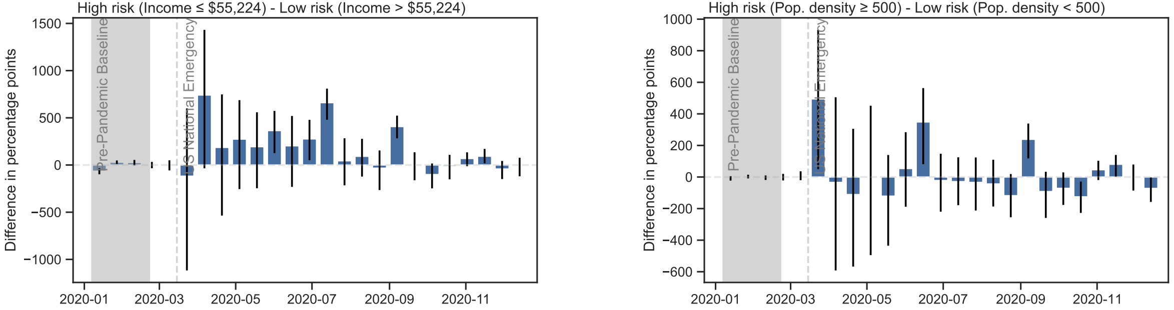
(a) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across 'Education level'.
High risk (Attained BA ≤ 21%) - Low risk (Attained BA > 21%)

(b) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across '% Insurance coverage'.
High risk (Insurance cov. ≤ 93%) - Low risk (Insurance cov. > 93%)



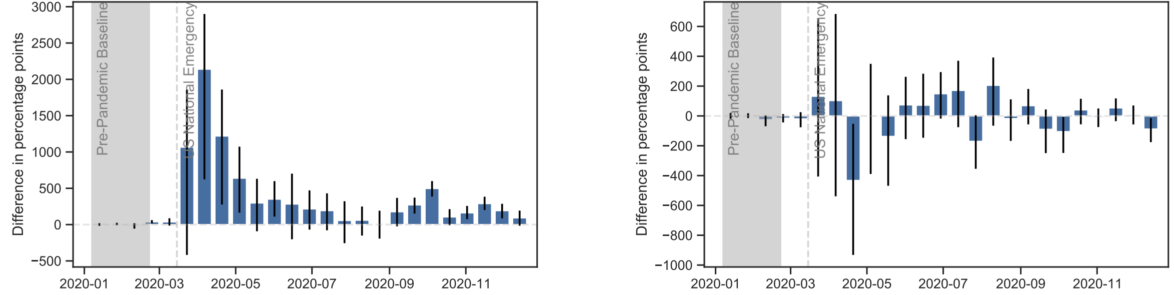
(c) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across '% Internet access'.
High risk (Internet access ≤ 82%) - Low risk (Internet access > 82%)

(d) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across '% Hispanic populations'.
High risk (Hisp. pop. ≥ 18%) - Low risk (Hisp. pop. < 18%)



(e) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across 'Income'.
High risk (Income ≤ $55,224) - Low risk (Income > $55,224)

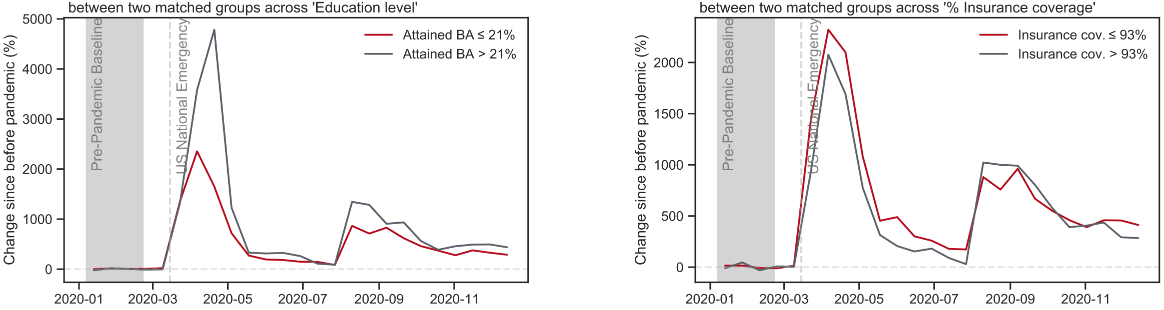
(f) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across 'Population density'.
High risk (Pop. density ≥ 500) - Low risk (Pop. density < 500)



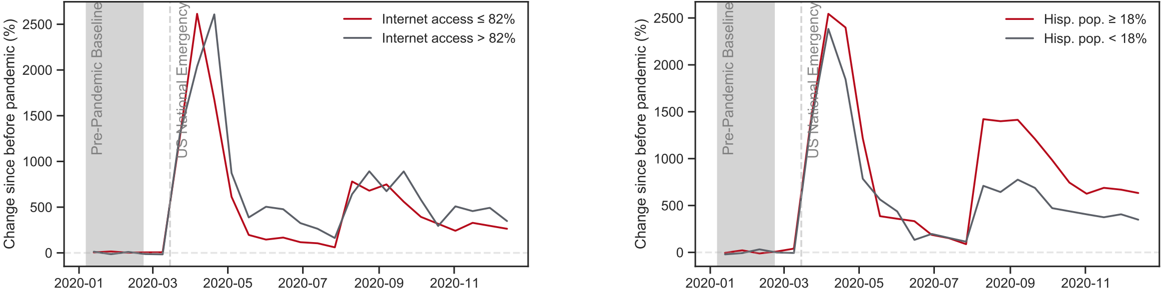
(g) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across '% Black population'.
High risk (Black pop. ≥ 12%) - Low risk (Black pop. < 12%)

(h) Diff. in percentage points for changes in 'Unemployment related queries' between two matched groups across '% Unemployed'.
High risk (Unemployed ≥ 3%) - Low risk (Unemployed < 3%)

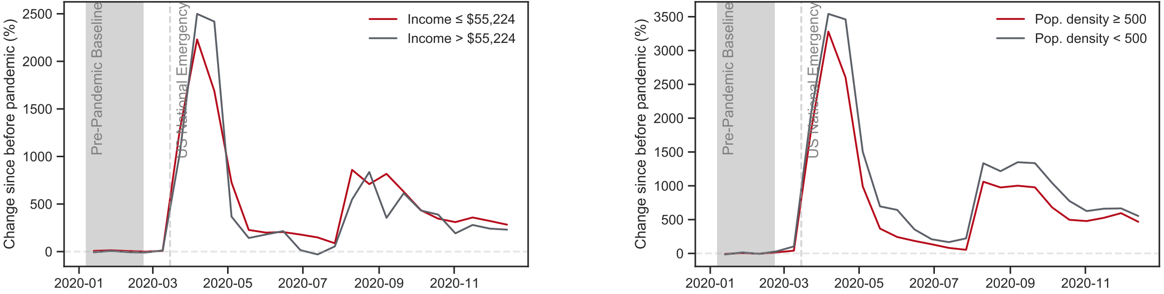
Figure B.17: Percent change in 'Clicks to state unemployment sites' between two matched groups across eight SDoH factors.


(a) Percent change in 'Clicks to state unemployment sites' between two matched groups across 'Education level'

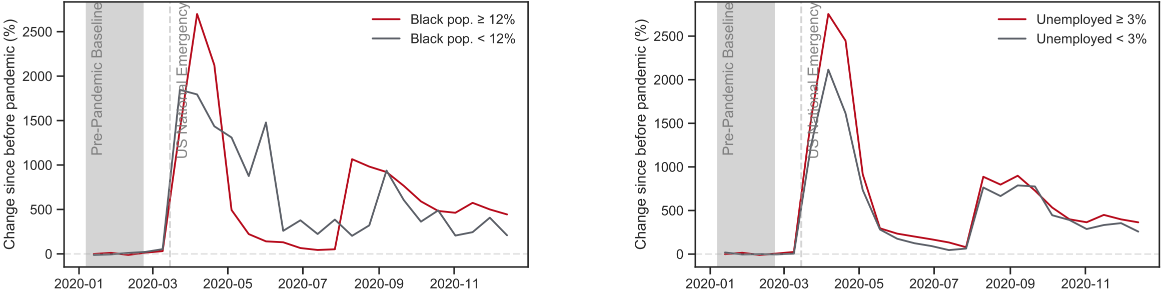
(b) Percent change in 'Clicks to state unemployment sites' between two matched groups across '% Insurance coverage'


(c) Percent change in 'Clicks to state unemployment sites' between two matched groups across '% Internet access'

(d) Percent change in 'Clicks to state unemployment sites' between two matched groups across '% Hispanic populations'


(e) Percent change in 'Clicks to state unemployment sites' between two matched groups across 'Income'

(f) Percent change in 'Clicks to state unemployment sites' between two matched groups across 'Population density'


(g) Percent change in 'Clicks to state unemployment sites' between two matched groups across '% Black population'

(h) Percent change in 'Clicks to state unemployment sites' between two matched groups across '% Unemployed'

Figure B.18: Differences in percentage points for changes in 'Clicks to state unemployment sites' between two matched groups across eight SDoH factors.

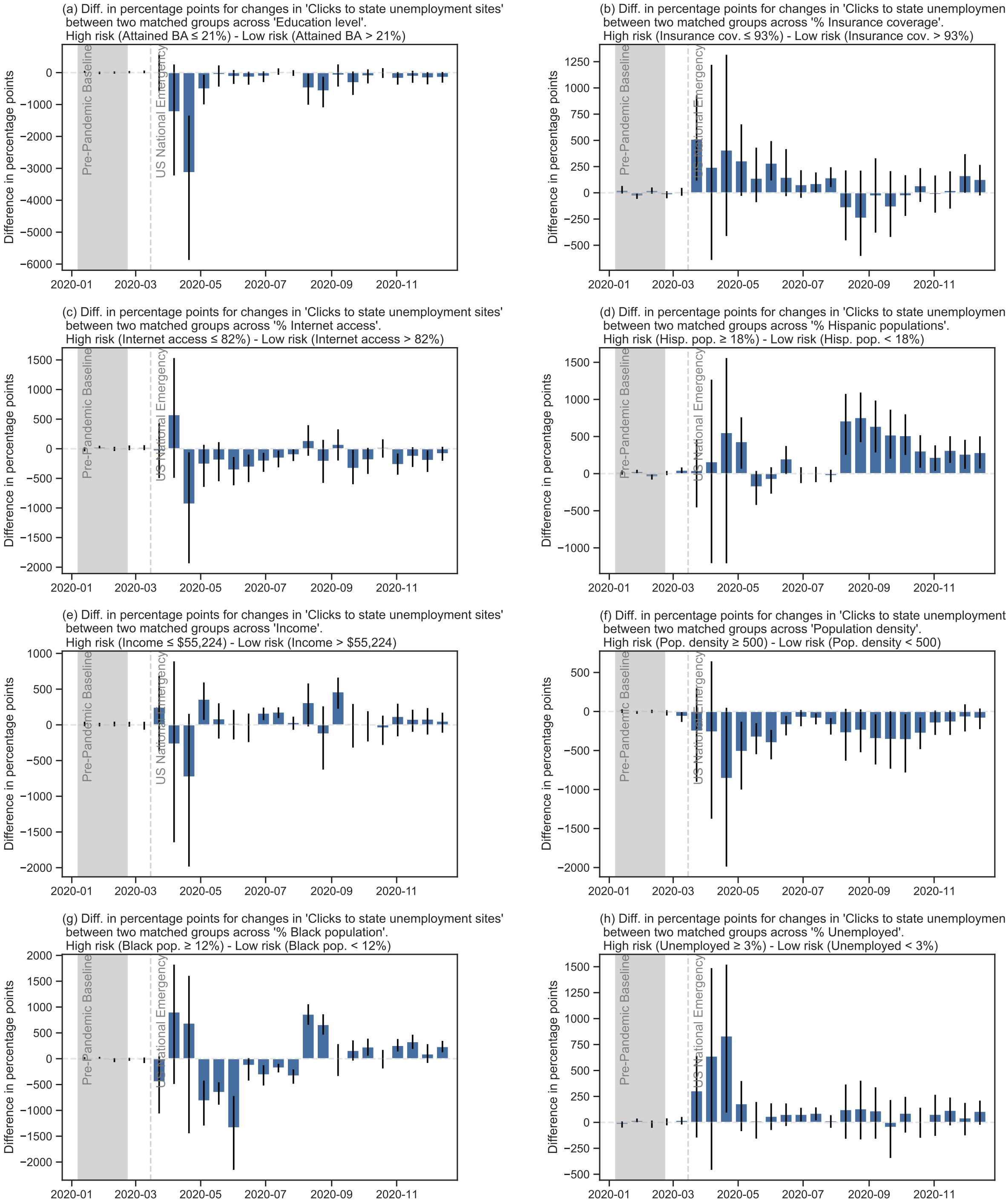
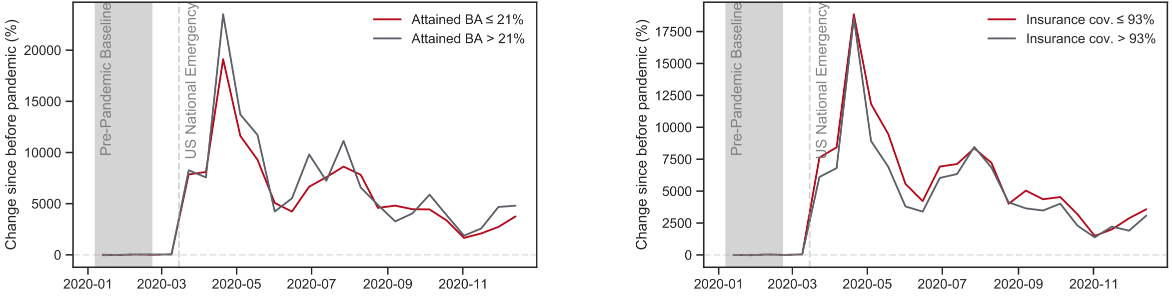
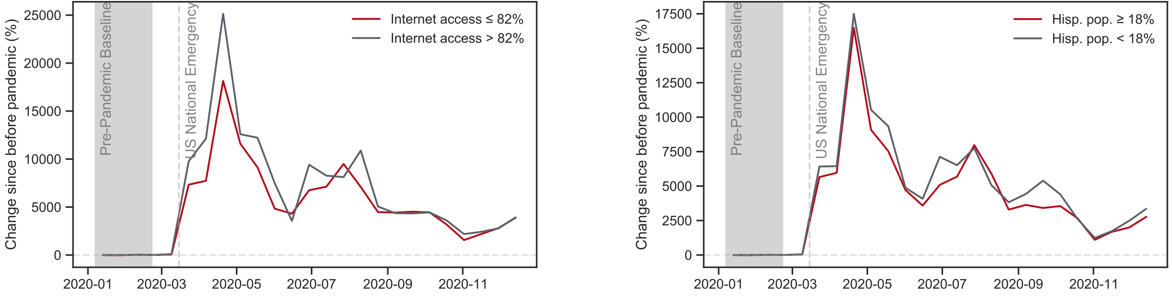


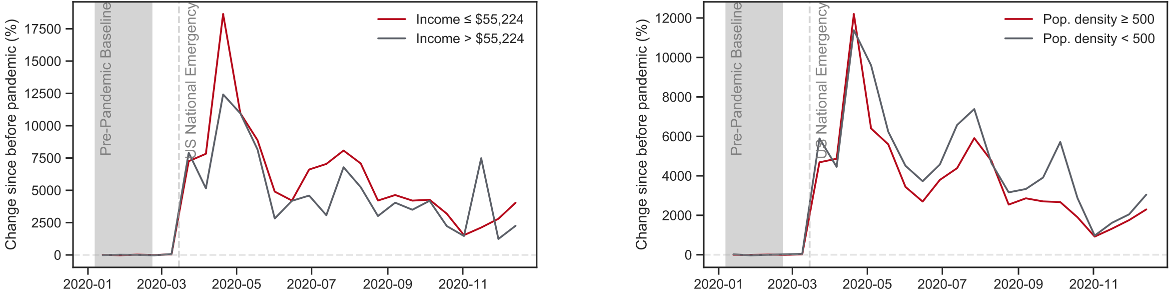
Figure B.19: Percent change in 'Financial assistance related queries' between two matched groups across eight SDoH factors.


(a) Percent change in 'Financial assistance related queries' between two matched groups across 'Education level'

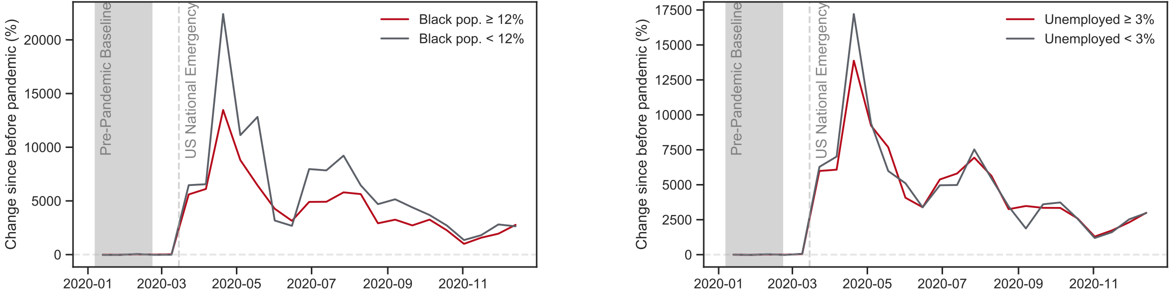
(b) Percent change in 'Financial assistance related queries' between two matched groups across '% Insurance coverage'


(c) Percent change in 'Financial assistance related queries' between two matched groups across '% Internet access'

(d) Percent change in 'Financial assistance related queries' between two matched groups across '% Hispanic populations'


(e) Percent change in 'Financial assistance related queries' between two matched groups across 'Income'

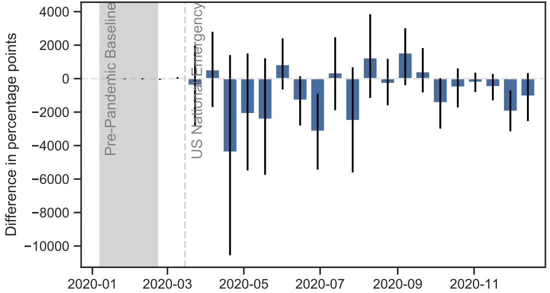
(f) Percent change in 'Financial assistance related queries' between two matched groups across 'Population density'


(g) Percent change in 'Financial assistance related queries' between two matched groups across '% Black population'

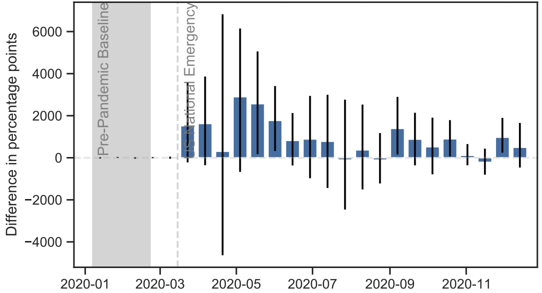
(h) Percent change in 'Financial assistance related queries' between two matched groups across '% Unemployed'

Figure B.20: Differences in percentage points for changes in ‘Financial assistance related queries’ between two matched groups across eight SDoH factors.

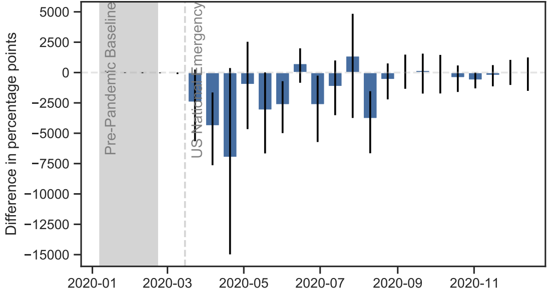
(a) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across 'Education level'.
High risk (Attained BA ≤ 21%) - Low risk (Attained BA > 21%)



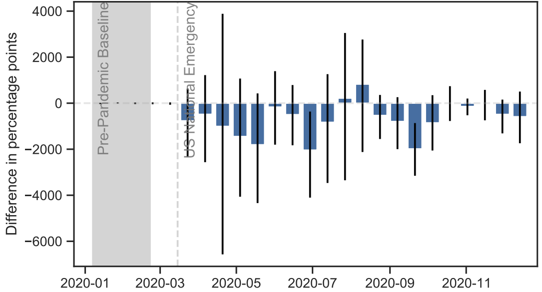
(b) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across '% Insurance coverage'.
High risk (Insurance cov. ≤ 93%) - Low risk (Insurance cov. > 93%)



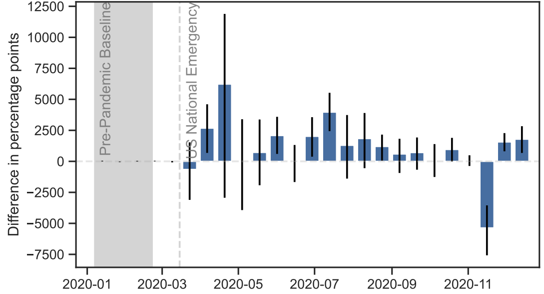
(c) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across '% Internet access'.
High risk (Internet access ≤ 82%) - Low risk (Internet access > 82%)



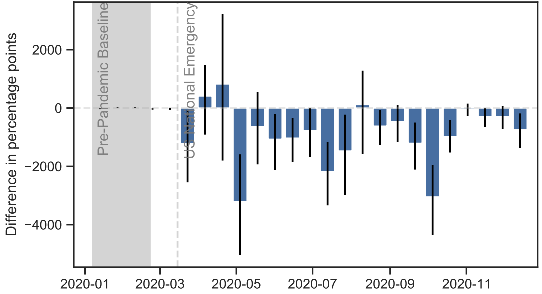
(d) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across '% Hispanic populations'.
High risk (Hisp. pop. ≥ 18%) - Low risk (Hisp. pop. < 18%)



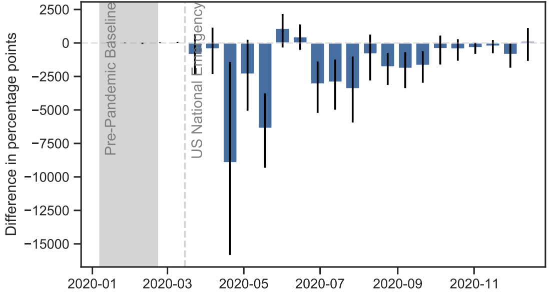
(e) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across 'Income'.
High risk (Income ≤ $55,224) - Low risk (Income > $55,224)



(f) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across 'Population density'.
High risk (Pop. density ≥ 500) - Low risk (Pop. density < 500)



(g) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across '% Black population'.
High risk (Black pop. ≥ 12%) - Low risk (Black pop. < 12%)



(h) Diff. in percentage points for changes in 'Financial assist. related queries' between two matched groups across '% Unemployed'.
High risk (Unemployed ≥ 3%) - Low risk (Unemployed < 3%)

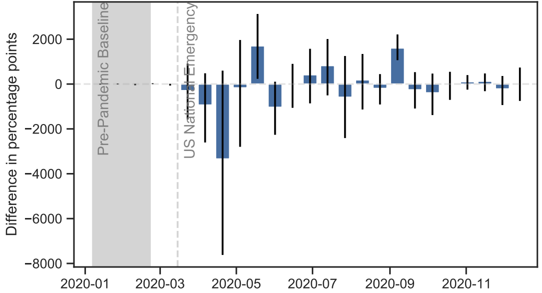
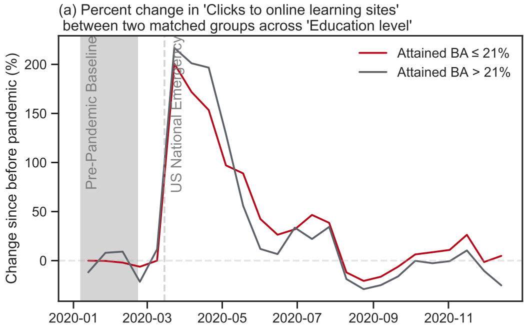
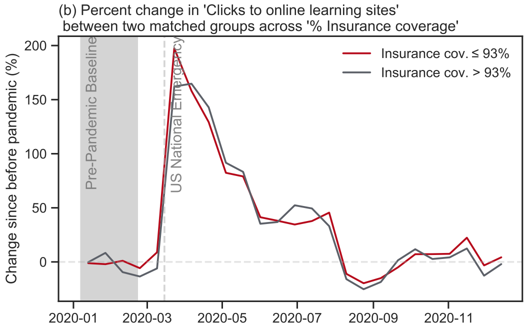
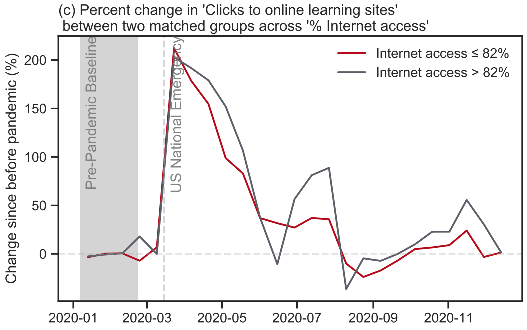
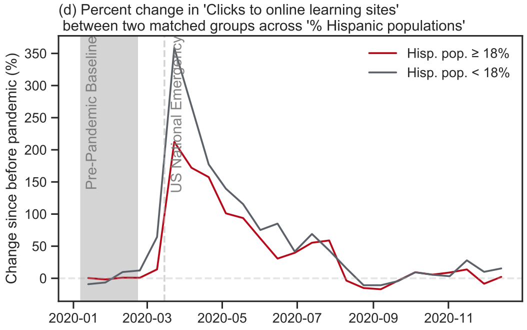


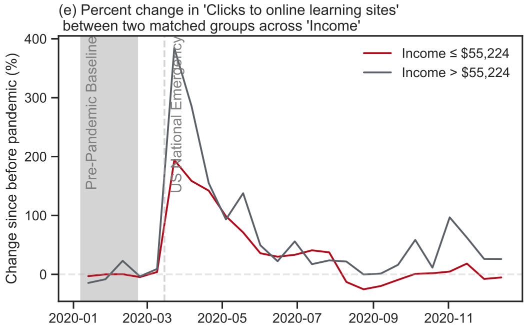
Figure B.21: Percent change in ‘Click to online learning sites’ between two matched groups across eight SDoH factors.

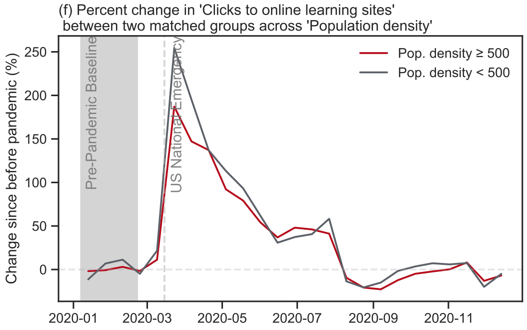

(a) Percent change in 'Clicks to online learning sites' between two matched groups across 'Education level'

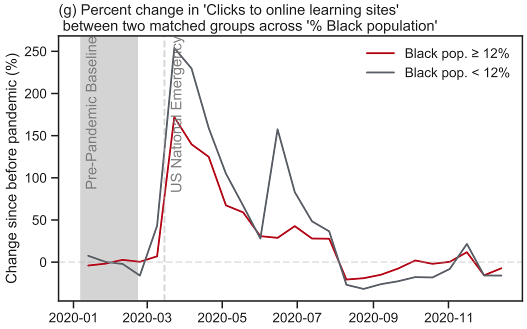

(b) Percent change in 'Clicks to online learning sites' between two matched groups across '% Insurance coverage'


(c) Percent change in 'Clicks to online learning sites' between two matched groups across '% Internet access'


(d) Percent change in 'Clicks to online learning sites' between two matched groups across '% Hispanic populations'


(e) Percent change in 'Clicks to online learning sites' between two matched groups across 'Income'


(f) Percent change in 'Clicks to online learning sites' between two matched groups across 'Population density'


(g) Percent change in 'Clicks to online learning sites' between two matched groups across '% Black population'

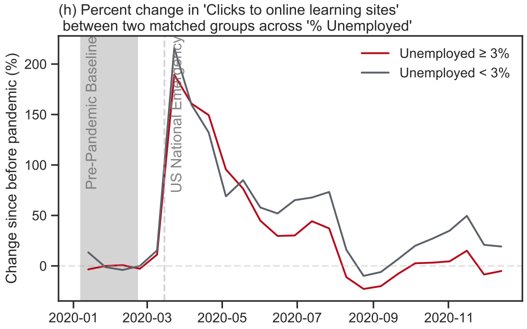
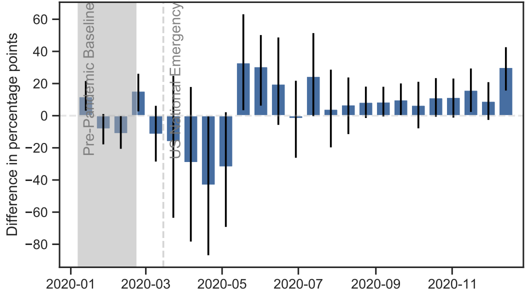
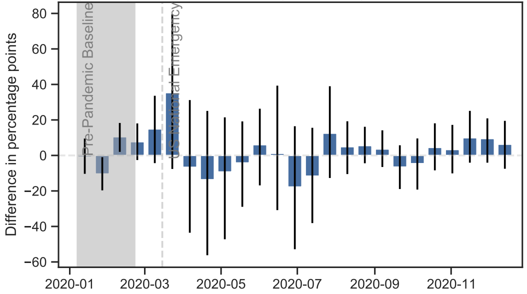

(h) Percent change in 'Clicks to online learning sites' between two matched groups across '% Unemployed'

Figure B.22: Differences in percentage points for changes in ‘Click to online learning sites’ between two matched groups across eight SDoH factors.

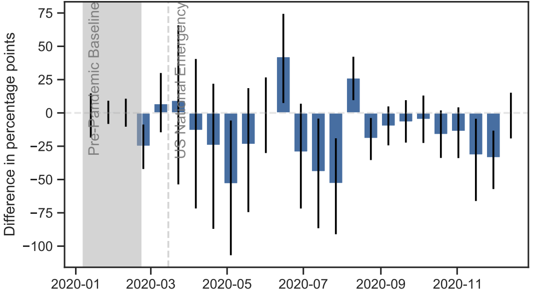
(a) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across 'Education level'.
High risk (Attained BA ≤ 21.1%) - Low risk (Attained BA > 21%)



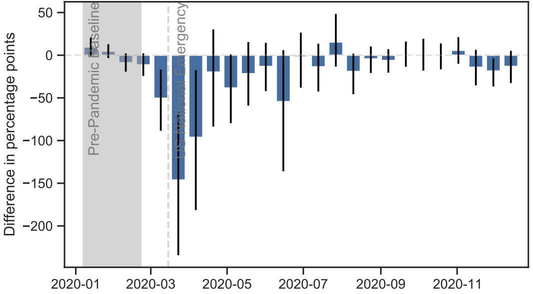
(b) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across '% Insurance coverage'.
High risk (Insurance cov. ≤ 92.7%) - Low risk (Insurance cov. > 92.7%)



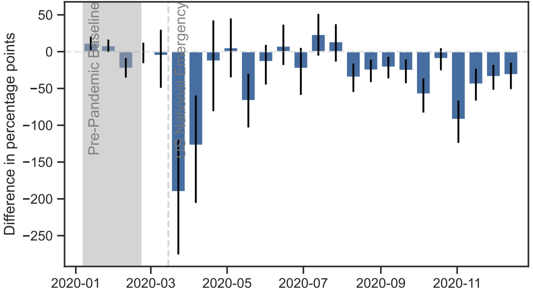
(c) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across '% Internet access'.
High risk (Internet access ≤ 81.8%) - Low risk (Internet access > 81.8%)



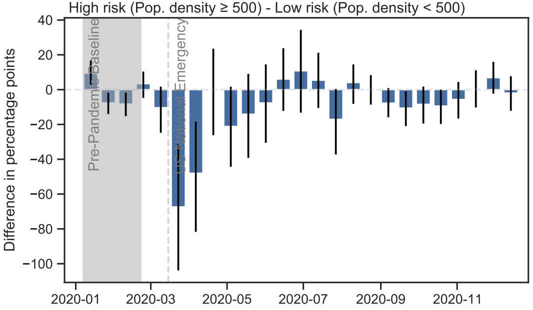
(d) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across '% Hispanic populations'.
High risk (Hisp. pop. ≥ 18%) - Low risk (Hisp. pop. < 18%)



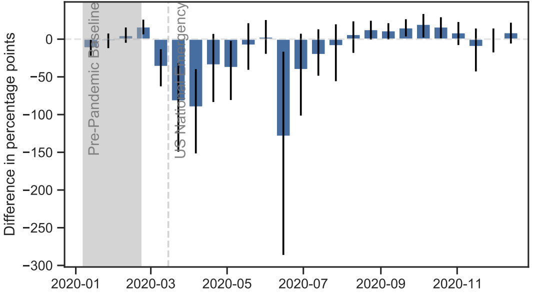
(e) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across 'Income'.
High risk (Income ≤ $55,224) - Low risk (Income > $55,224)



(f) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across 'Population density'.
High risk (Pop. density ≥ 500) - Low risk (Pop. density < 500)



(g) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across '% Black population'.
High risk (Black pop. ≥ 12%) - Low risk (Black pop. < 12%)



(h) Diff. in percentage points for changes in 'Clicks to online learning sites'
between two matched groups across '% Unemployed'.
High risk (Unemployed ≥ 3%) - Low risk (Unemployed < 3%)

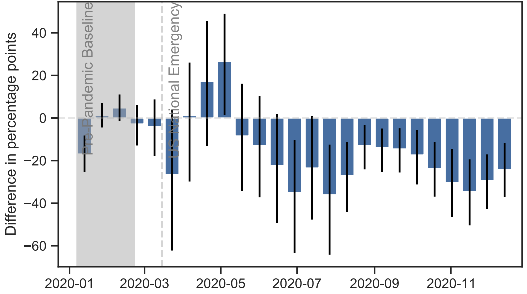
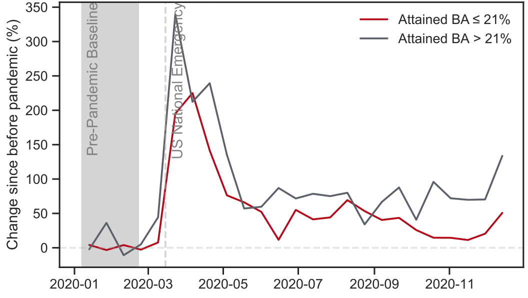
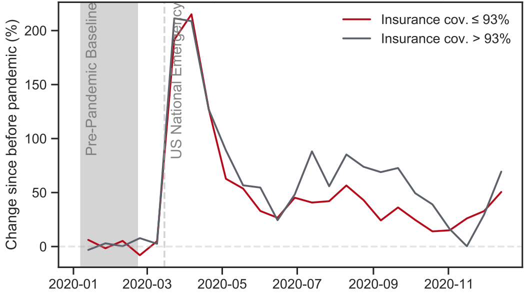


Figure B.23: Percent change in ‘Food delivery related queries’ between two matched groups across eight SDoH factors.

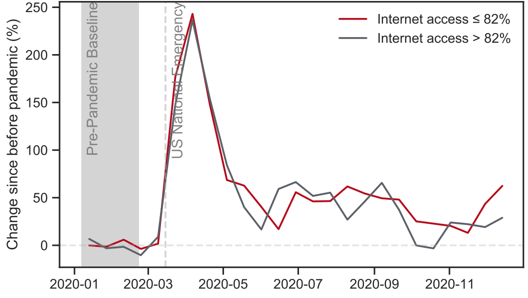
(a) Percent change in 'Food delivery related queries' between two matched groups across 'Education level'



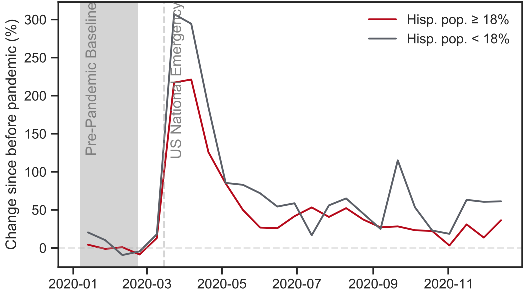
(b) Percent change in 'Food delivery related queries' between two matched groups across '% Insurance coverage'



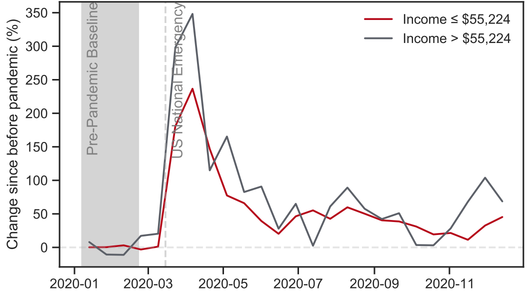
(c) Percent change in 'Food delivery related queries' between two matched groups across '% Internet access'



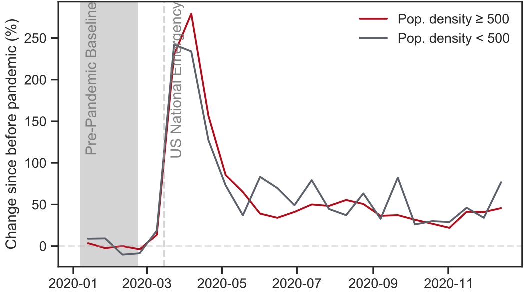
(d) Percent change in 'Food delivery related queries' between two matched groups across '% Hispanic populations'



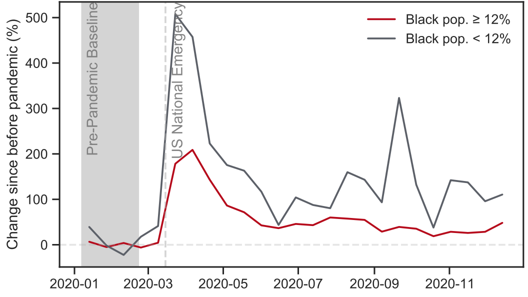
(e) Percent change in 'Food delivery related queries' between two matched groups across 'Income'



(f) Percent change in 'Food delivery related queries' between two matched groups across 'Population density'



(g) Percent change in 'Food delivery related queries' between two matched groups across '% Black population'



(h) Percent change in 'Food delivery related queries' between two matched groups across '% Unemployed'

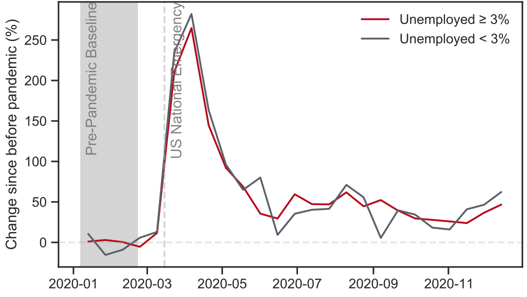
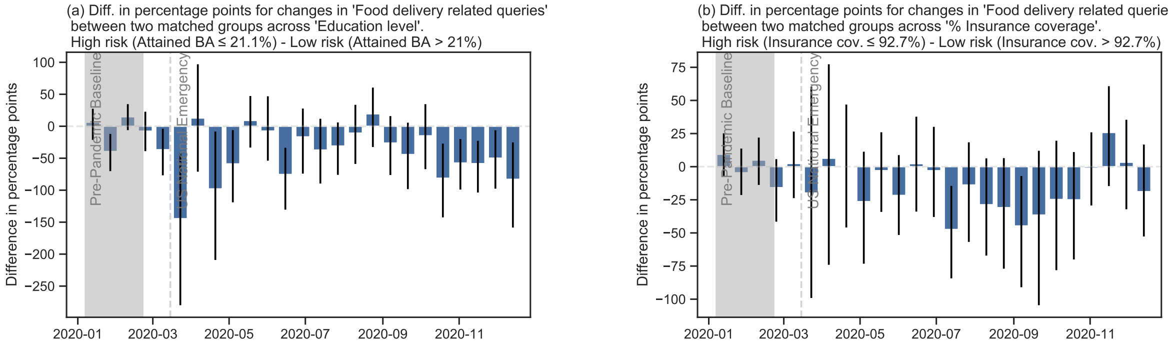
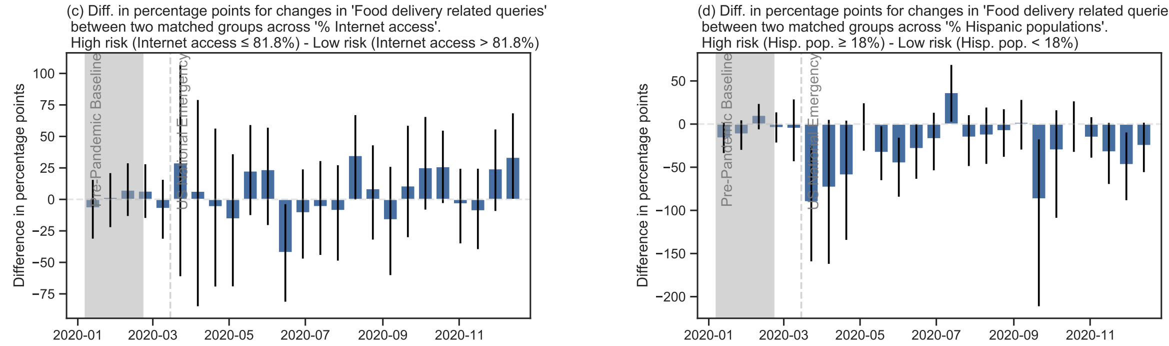


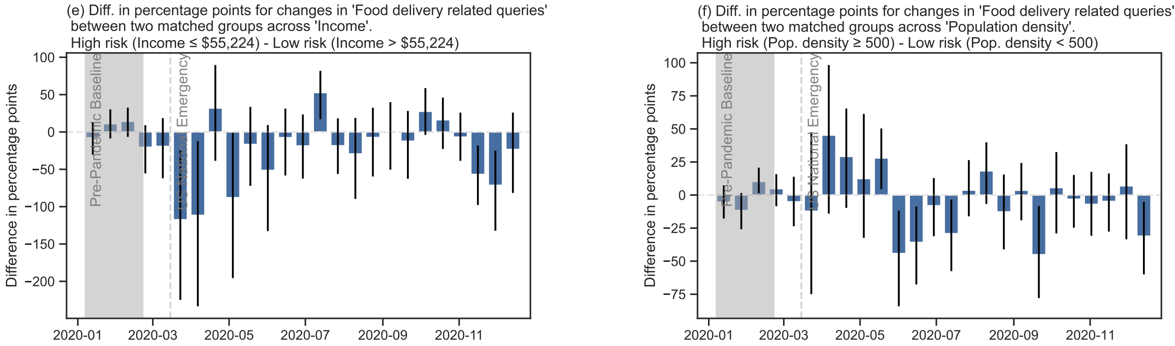
Figure B.24: Differences in percentage points for changes in ‘Food delivery related queries’ between two matched groups across eight census variables.


(a) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across 'Education level'. High risk (Attained BA ≤ 21.1%) - Low risk (Attained BA > 21%)

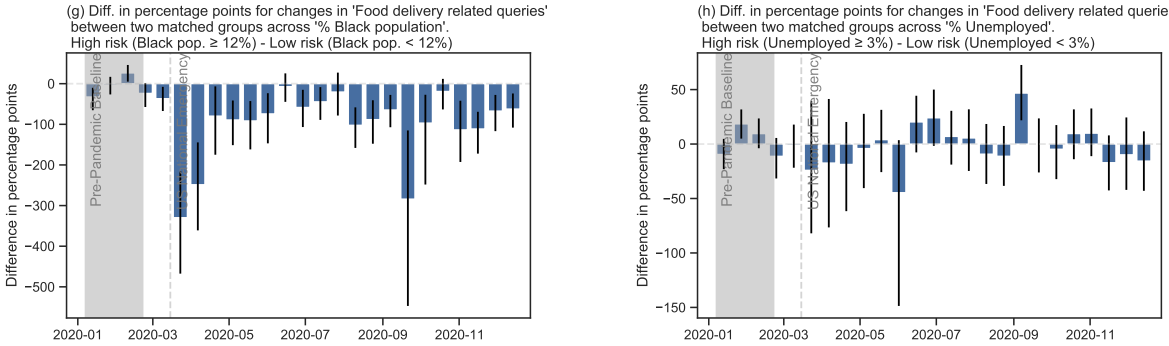
(b) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across '% Insurance coverage'. High risk (Insurance cov. ≤ 92.7%) - Low risk (Insurance cov. > 92.7%)


(c) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across '% Internet access'. High risk (Internet access ≤ 81.8%) - Low risk (Internet access > 81.8%)

(d) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across '% Hispanic populations'. High risk (Hisp. pop. ≥ 18%) - Low risk (Hisp. pop. < 18%)


(e) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across 'Income'. High risk (Income ≤ $55,224) - Low risk (Income > $55,224)

(f) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across 'Population density'. High risk (Pop. density ≥ 500) - Low risk (Pop. density < 500)

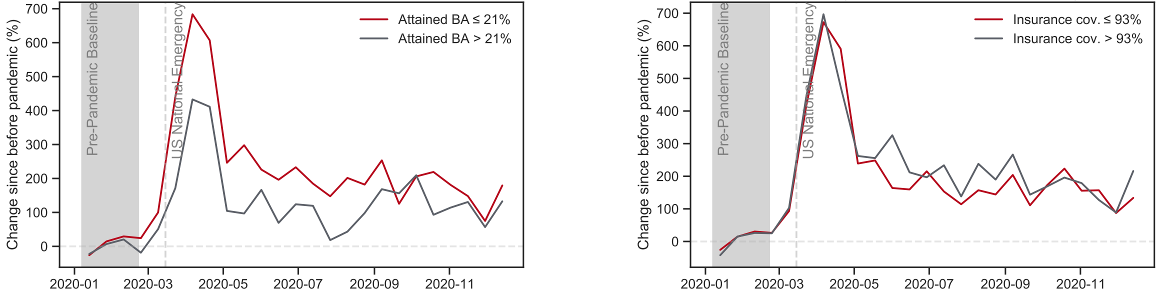

(g) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across '% Black population'. High risk (Black pop. ≥ 12%) - Low risk (Black pop. < 12%)

(h) Diff. in percentage points for changes in 'Food delivery related queries' between two matched groups across '% Unemployed'. High risk (Unemployed ≥ 3%) - Low risk (Unemployed < 3%)

Figure B.25: Percent change in ‘Food assistance related queries’ between two matched groups across eight SDoH factors.

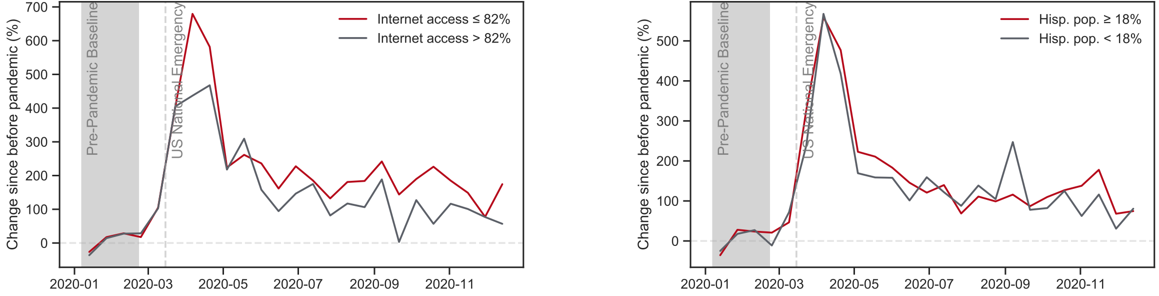
(a) Percent change in 'Food assistance related queries' between two matched groups across 'Education level'

(b) Percent change in 'Food assistance related queries' between two matched groups across '% Insurance coverage'



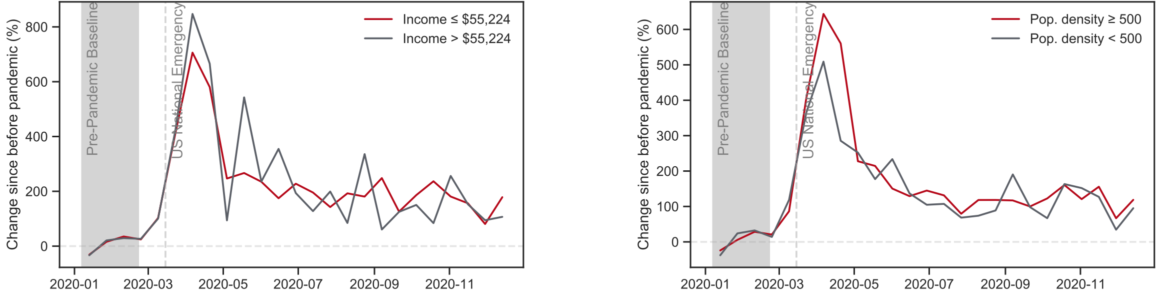
(c) Percent change in 'Food assistance related queries' between two matched groups across '% Internet access'

(d) Percent change in 'Food assistance related queries' between two matched groups across '% Hispanic populations'



(e) Percent change in 'Food assistance related queries' between two matched groups across 'Income'

(f) Percent change in 'Food assistance related queries' between two matched groups across 'Population density'



(g) Percent change in 'Food assistance related queries' between two matched groups across '% Black population'

(h) Percent change in 'Food assistance related queries' between two matched groups across '% Unemployed'

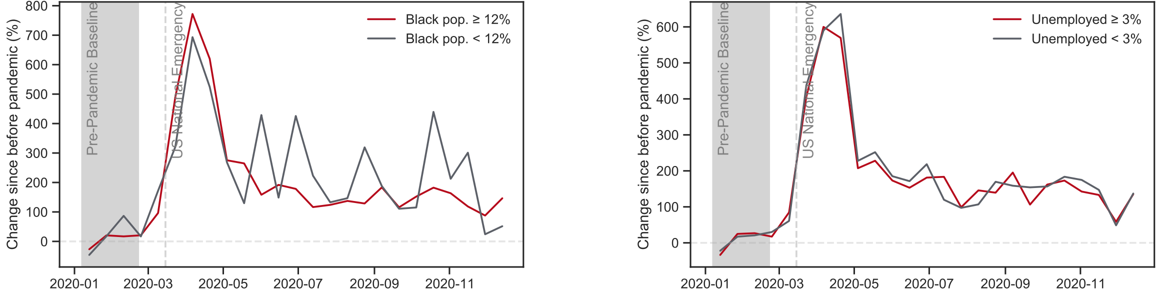
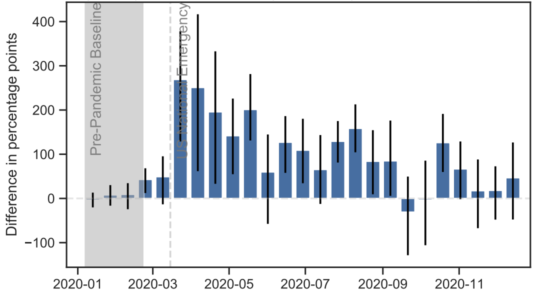
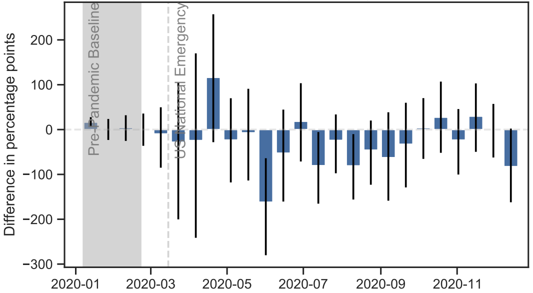
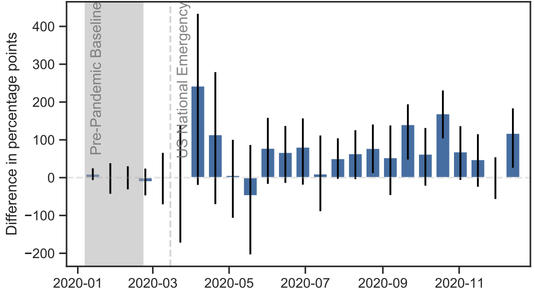
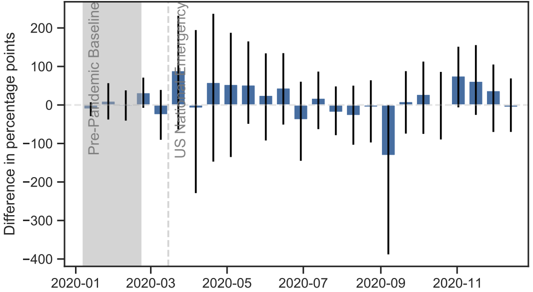


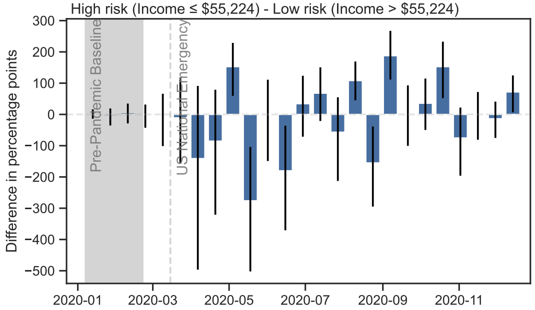
Figure B.26: Differences in percentage points for changes in 'Food assistance related queries' between two matched groups across eight SDoH factors.

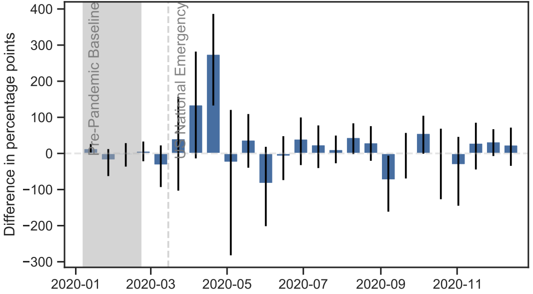

(a) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across 'Education level'.
High risk (Attained BA ≤ 21.1%) - Low risk (Attained BA > 21%)

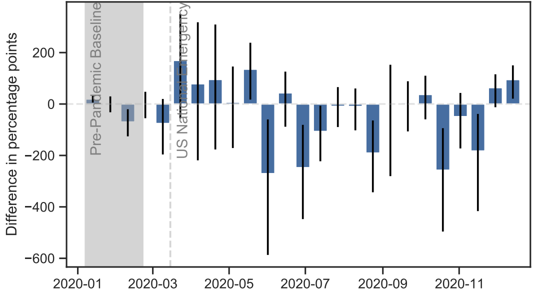

(b) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across '% Insurance coverage'.
High risk (Insurance cov. ≤ 92.7%) - Low risk (Insurance cov. > 92.7%)

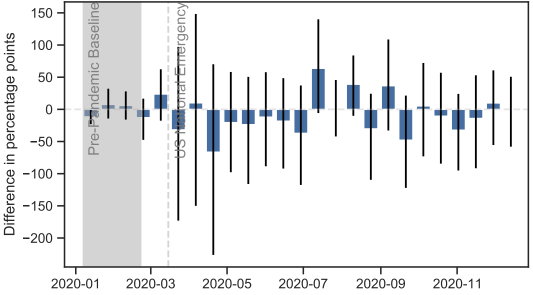

(c) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across '% Internet access'.
High risk (Internet access ≤ 81.8%) - Low risk (Internet access > 81.8%)


(d) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across '% Hispanic populations'.
High risk (Hisp. pop. ≥ 18%) - Low risk (Hisp. pop. < 18%)


(e) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across 'Income'.
High risk (Income ≤ $55,224) - Low risk (Income > $55,224)


(f) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across 'Population density'.
High risk (Pop. density ≥ 500) - Low risk (Pop. density < 500)


(g) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across '% Black population'.
High risk (Black pop. ≥ 12%) - Low risk (Black pop. < 12%)


(h) Diff. in percentage points for changes in 'Food assistance related queries' between two matched groups across '% Unemployed'.
High risk (Unemployed ≥ 3%) - Low risk (Unemployed < 3%)

Appendix C

QUESTIONNAIRES FOR STRESS REDUCTION INTERVENTION STUDY

C.1 ECOLOGICAL MOMENTARY ASSESSMENTS (EMAS)

EMA questions were presented to participants five times throughout the day. The timing of these questions was determined by each participant's self-reported start and end of the day. The first EMA was scheduled at the start of the workday, and the remaining EMAs were slotted to each of the remaining 4 intervals at the half hour or hour marks (e.g., 2:00 or 2:30), ensuring that there is no more than one EMA per hour and that the last EMA is done at least 30 minutes before the end of the workday. For example, if a participant started their workday at 9 AM and concluded their workday at 5 PM, we presented EMAs at 9 AM, 11 AM, 12:30 PM, 2:30 PM, and 4 PM.

Below are the questions that were presented in these EMAs:

1. How would you rate your level of stress during the last 30 minutes? *Not at all – Slightly stressed – Moderately stressed – Very stressed – Extremely stressed*
2. Considering the last 30 minutes, how would you rate the demands on you? *very low – low – moderate – high – very high*
3. Considering the last 30 minutes, how would you rate the resources you had available? *very low – low – moderate – high – very high*
4. Considering the last 30 minutes, how would you rate the level of energy you experienced? *very low – low – moderate – high – very high*
5. Considering the last 30 minutes, how would you rate the level of pleasantness you

experienced? *very unpleasant – unpleasant – neutral – pleasant – very pleasant*

6. Considering the last 30 minutes, did you have any eating episodes? *No – Yes*
7. Considering the last 30 minutes, how many social interactions did you have? *0 – 1-2 – 3-4 – >4*

C.2 POST-STUDY QUESTIONNAIRE

All study participants were asked to answer post-study questions. There were questions common to both conditions as well as questions specific to each condition. Some condition-specific questions were mirrored across the conditions to ensure that each question was trying to get at the same concept but about the system design presented in each condition.

Below are the questions that were presented in these post-study questionnaires for each condition, including common questions. Post-study questionnaire also included DASS-21 in addition to the questions below.

C.2.1 Pre-scheduled condition

During the study, you were instructed to schedule interventions for the upcoming week in your calendar by browsing the intervention catalog. Remember, the term "interventions" refers to the brief stress-reducing activities Huey the bot delivered to you – activities like deep breathing or calling a friend for support.

We fully acknowledge that there were performance and functional bugs with the bot and the sensing software. Putting the obvious bugs aside, we would like you to think about the potential for this intervention technology as you answer the following questions.

Please rate your agreement with the following statements. *Strongly disagree – Somewhat disagree – Neither agree nor disagree – Somewhat agree – Strongly agree*

1. The intervention bot meets my requirements for engaging in interventions.
2. Using the intervention bot is a frustrating experience.
3. The intervention bot is easy to use.
4. The intervention bot is helpful for reducing my workplace stress.
5. I found myself engaging in more interventions with the intervention bot (compared to before the study).
6. I found that the intervention bot made it easier to engage in interventions (compared to before the study).
7. I am satisfied with the intervention bot.
8. I would continue to use the intervention bot.

Please answer the following open-ended questions as best as you can to help us improve the user experience.

Remember, please don't include any personally identifying information in your responses below, such as specific names or locations.

1. What are some factors that you took into account when scheduling your interventions (e.g., time of day, day of week, specific intervention activity, etc.)?
2. What did you like about scheduling interventions in advance?
3. What did you dislike about scheduling interventions in advance?
4. If you performed interventions on demand, how did performing on-demand interventions compare to scheduling them in advance?
5. How would you compare proactively scheduling times to do interventions (as you have done in the study) to a system nudging you to do interventions when it detects you need them to reduce and manage stress or improve your wellbeing? Which would you prefer and why?
6. How did you find the timing of the interventions that you scheduled in advance? Did you find that you were able to know in advance when a good time for doing interventions would be?

7. Do you have any suggestions for improving the experience of scheduling interventions in advance?

C.2.2 Just-in-time condition

During the study, you were instructed to perform interventions when the bot nudged you to do them. Remember, the term “interventions” refers to the brief stress-reducing activities Huey the bot delivered to you – activities like deep breathing or calling a friend for support.

We fully acknowledge that there were performance and functional bugs with the bot and the sensing software. Putting the obvious bugs aside, we would like you to think about the potential for this intervention technology as you answer the following questions.

Please rate your agreement with the following statements. *Strongly disagree – Somewhat disagree – Neither agree nor disagree – Somewhat agree – Strongly agree*

1. The intervention bot meets my requirements for engaging in interventions.
2. Using the intervention bot is a frustrating experience.
3. The intervention bot is easy to use.
4. The intervention bot is helpful for reducing my workplace stress.
5. I found myself engaging in more interventions with the intervention bot (compared to before the study).
6. I found that the intervention bot made it easier to engage in interventions (compared to before the study).
7. I am satisfied with the intervention bot.
8. I would continue to use the intervention bot.

Please answer the following open-ended questions as best as you can to help us improve the user experience.

Remember, please don't include any personally identifying information in your responses below, such as specific names or locations.

1. What did you like about being nudged to do interventions?
2. What did you dislike about being nudged to do interventions?
3. If you performed interventions on demand, how did performing on-demand interventions compare to being nudged to do interventions?
4. How would you compare being nudged to do interventions (as you have in this study) to proactively scheduling times to do interventions on your calendar? Which would you prefer and why?
5. How did you find the timing of these nudges to do interventions? Did you find that the nudges were sent at appropriate times?
6. Do you have any suggestions for improving the experience of being nudged to do interventions?

C.2.3 Common

Please answer the following open-ended questions as best as you can to help us improve the user experience.

Remember, please don't include any personally identifying information in your responses below, such as specific names or locations.

1. What kind of interventions were the most helpful in immediately reducing your stress levels?
2. What kind of interventions had a lasting positive impact on your stress levels (e.g., reduced stress for the rest of the day or even the day after)?
3. What kind of interventions did you like the most?
4. What kind of interventions were the least helpful in immediately reducing your stress levels?

5. What kind of interventions had a lasting negative impact on your stress levels, e.g. increased stress for the rest of the day or even the day after?
6. What kind of interventions did you like the least?
7. Based on your experience of performing the interventions during the study, what factors most likely influence your decision to do the interventions? Some examples include timing, simplicity, types of interventions, etc. Please elaborate as best as you can.
8. How did you find the interventions being facilitated by a Teams bot?
9. Do you have any feedback for improving the intervention design? Any feedback (content, quality, modality, etc.) are welcome.
10. By participating in this study, what changed for you? Did you actually spend more time taking care of yourself? Did you develop any new routines?
11. How did you feel about the check-ins required by the study? Check-ins were questions asking for your current stress level 4-5 times throughout the workday and twice per day on the weekends.
12. Do you have any feedback for improving the bot design? Any feedback (functionality, persona, modality, etc.) is welcome.
13. Do you have any feedback for the sensing software?
14. Please share any feedback or comments about your overall experience with the systems presented in the study, including ideas about the future of technology that could assist in planning and engaging in interventions and personalizing interventions that work for you.
15. Please share any feedback or comments about your overall study participation experience.

Listed below are a number of difficult or stressful things that sometimes happen to people.

- Natural disaster (for example, flood, hurricane, tornado, earthquake)

- Fire or explosion
- Transportation accident (for example, car accident, boat accident, train wreck, plane crash)
- Serious accident at work, home, or during recreational activity
- Exposure to toxic substance (for example, dangerous chemicals, radiation)
- Physical assault (for example, being attacked, hit, slapped, kicked, beaten up)
- Assault with a weapon (for example, being shot, stabbed, threatened with a knife, gun, bomb)
- Sexual assault (rape, attempted rape, made to perform any type of sexual act through force or threat of harm)
- Other unwanted or uncomfortable sexual experience
- Combat or exposure to a war-zone (in the military or as a civilian)
- Captivity (for example, being kidnapped, abducted, held hostage, prisoner of war)
- Life-threatening illness or injury
- Severe human suffering
- Sudden violent death (for example, homicide, suicide)
- Sudden accidental death
- Serious injury, harm, or death you caused to someone else
- Any other very stressful event or experience

Please indicate how many times ANY of the above events you have experienced in the past three months. *Never – 1-3 times – 4-6 times – 7+ times – Prefer not to answer*

How strongly do you agree or disagree with the following statements? *Strongly disagree – Somewhat disagree – Neither agree nor disagree – Somewhat agree – Strongly agree*

1. I tend to bounce back quickly after hard times.
2. I have a hard time making it through stressful events.

3. It does not take me long to recover from a stressful event.
4. It is hard for me to snap back when something bad happens.
5. I usually come through difficult times with little trouble.
6. I tend to take a long time to get over setbacks in my life.

Based on your engagement with this study, we assume you'd like to reduce your stress at work. Which stage are you currently in regarding changing your behavior to reduce stress at work *Precontemplation: Not ready to engage (zero prior engagement) – Contemplation: Getting ready to engage (engaged but zero action) – Taking action: Engaging but everything is still new – Maintenance: Changed behavior 6+ months ago*