Understanding Client Support Strategies to Improve Clinical Outcomes in an Online Mental Health Intervention

Prerna Chikersal[1], Danielle Belgrave[2], Gavin Doherty[3], Angel Enrique[4], Jorge E. Palacios[4], Derek Richards[4] and Anja Thieme[2]

[1]Carnegie Mellon University, prerna@cmu.edu

[2]Microsoft Research Cambridge, (dabelgra, anthie)@microsoft.com

[3]Trinity College Dublin, gavin.doherty@tcd.ie

[4]SilverCloud Health, (angel.enrique, jorge.palacios, derek.richards)@silvercloudhealth.com

ABSTRACT

Online mental health interventions are increasingly important in providing access to, and supporting the effectiveness of, mental health treatment. While these technologies are effective, user attrition and early disengagement are key challenges. Evidence suggests that integrating a human supporter into such services mitigates these challenges, however, it remains under-studied *how* supporter involvement benefits client outcomes, and how to maximize such effects. We present our analysis of 234,735 supporter messages to discover how different support strategies correlate with clinical outcomes. We describe our machine learning methods for: (i) clustering supporters based on client outcomes; (ii) extracting and analyzing linguistic features from supporter messages; and (iii) identifying context-specific patterns of support. Our findings indicate that *concrete, positive* and *supportive* feedback from supporters that reference *social behaviors* are strongly associated with better outcomes; and show how their importance varies dependent on different client situations. We discuss design implications for personalized support and supporter interfaces.

Author Keywords

Mental health; digital behavioral intervention; CBT; support; machine learning; AI; unsupervised learning; data mining.

CCS Concepts

•Human-centered computing → Human computer interaction; •Computing methodologies → *Machine learning;*

INTRODUCTION

Mental illness is increasing in occurrence [39]. It presents the largest cause of disability worldwide and is the strongest predictor of suicide [64, 99]. This makes the prevention and treatment of mental health disorders a public health priority [102] and has led to explorations of how the field of HCI, and the development of technology more broadly, can support access to, and increase the effectiveness of, mental health treatment [8, 76, 95, 86]. Over the last decade, this has brought forward developments of mobile apps [19, 29, 55], and computerized psycho-educational and psycho-therapeutic interventions [5, 17, 81, 103], or chat-based [27, 48, 88] programs to complement, and expand access to, psychotherapy.

Most existing digital mental health services are based on Cognitive Behavioral Therapy (CBT); the most widely applied and most extensively empirically tested psychotherapy in Western Healthcare [11]. CBT is solution-focused, teaches the person to attend to the relationships between their thoughts, feelings and behaviors, and is frequently used in treating depression, anxiety or post-traumatic stress. Its highly structured format makes it well suited for support by digital technology [18]. Further, extensive research has evidenced the clinical effectiveness of internet-delivered CBT (iCBT) with sustainable results comparable to face-to-face therapy [4, 5, 100, 103].

Despite these benefits, a key challenge for digital behavioral interventions like iCBT is sustaining the users' engagement with treatment [26, 43], where early disengagement and dropout from the therapy program can mean users may not get the desired benefits. Thus, approaches to design an engaging online experience to sustain use and ensure beneficial health outcomes have become a deliberate focus for (HCI) research and development. This often means increasing opportunities for: (i) *interactivity*; (ii) *personalized experiences*; and (iii) *social support* through a community of (peer) moderators, trained supporters, or remote therapists [21, 50, 70, 79].

To aid engagement with online therapy, the involvement of a human supporter (*e.g.*, via text messages) has especially been shown to lead to more effective outcomes than unsupported interventions [40, 89, 103]. However, existing research on the effectiveness of supported-interventions has primarily assessed the impact of support duration and frequency [32, 45, 61, 96]; and to a lesser extent, different types of supporter behaviors [36, 69, 87]. Thus, it is less well understood how supporter behaviors impact program use, and therapeutic outcomes; and how this may differ between clients. Having a more nuanced understanding of the impact of supporter behaviors on clients however could help to: better maximize the

CHI '20, April 25–30, 2020, Honolulu, HI, USA

ISBN 978-1-4503-6708-0/20/04.

DOI: https://doi.org/10.1145/3313831.3376341

effect and outcomes of supporter involvement (*e.g.*, through more targeted and tailored forms of support), assist in supporter training, and thus, increase quality of care for clients.

Simultaneously, the rise in the uptake of internet-delivered therapies and increase in the scale of automatically collected usage data from these treatment programs enables new methodological possibilities for improving our understanding of human behaviors and optimizing health outcomes (*e.g.*, [6, 42, 56]). Specifically, the fields of data mining and machine learning (ML) provide advanced computational methods to construct robust systems that can automatically learn from, and identify important patterns in, behavioral data [91]. These techniques have been successfully used in gaming and for recommender systems; and show great potential for advancing our understanding of health data [28]. Thus, they are being increasingly explored in the context of mental health (*e.g.* [25, 34, 74]).

Our work presents the first application of unsupervised machine learning, and statistical and data mining methods for analyzing complex, large-scale supporter-client interaction data to identify supporter behaviors that correlate with better clinical outcomes. Our analysis is based on a fully anonymized dataset of 234,735 supporter messages to clients (sent by 3,481 supporters to 54,104 clients) from an established iCBT program for depression and anxiety, on the SilverCloud platform (www.silvercloudhealth.com), that delivers treatment with regular feedback messages sent by a human supporter. More specifically, our work makes the following contributions:

1. We describe our approach and the challenges involved in developing computational methods for this analysis. This includes: (i) *clustering supporters* based on how the support messages they sent to clients correlate with client outcomes; (ii) *extracting linguistic features* in support messages indicative of supporter behaviors that correlate with "high" outcomes across clients in different contexts or situations (*e.g.* different levels of usage); and (iii) taking into account co-occurrent patterns of different context variables and individual support strategies, we leverage data mining to identify salient *context-specific patterns* of support.

2. Our work indicates that *concrete, positive*, and *supportive* messages from supporters that reference *social behaviors* are strongly associated with better outcomes; and that the importance of support strategies can vary dependent on a clients' specific context (*e.g.* their mental health, platform use). Based on these findings, we discuss: (i) design implications for personalized support in iCBT interventions; (ii) the need for human-centeredness in health data science; and (iii) ethical considerations for secondary data analysis.

RELATED WORK

We begin with an overview of relevant literature that describes existing research on: (i) human support in digital mental health interventions; and (ii) data mining and ML approaches for understanding large-scale mental health behavior data.

Human Support in Online Mental Health Therapy

In online mental health interventions, the role of supporters, who can be trainees or therapists, often differs from the responsibilities of a therapist in more traditional face-to-face therapy. While supporters encourage and facilitate client use of an iCBT program, the clients themselves have to learn the necessary self-management knowledge and skills that are collectively the active ingredients of the intervention [81, 89].

Modalities & Benefit of Human Support in iCBT

Human support in digital mental health interventions can take various forms, ranging from different communication modes (*e.g.* email, text, phone, or video chat [16, 51]), to variations in support frequency and duration [32, 45, 61, 96], and support behaviors [36, 69, 87]. Most studies on the effects of human supported iCBT apps, programs or platforms, assess the therapeutic or working *alliance* —a bond that facilitates collaborative engagement in therapy [9] —between supporters and clients. The research suggests that such an alliance can be formed within iCBT environments [65] with consistent evidence of the benefits of support in those interventions [40, 82, 89, 103]. For example, Richards & Timulak [82] studied what clients identified as helpful or hindering in their treatment, and found that clients rated the helpfulness of supporter behaviors equal to the core content of the intervention. The literature however is less conclusive on how differences in communication medium [51], frequency [16] and duration [96] of support impact outcomes. For example, Titov [96] found no difference between interventions with low-intensity support (<3 hours) and high-intensity support (>3 hours).

Human Support Behaviors & their Effectiveness in iCBT

To date, only a small number of works have explicitly studied iCBT support behaviors and their impact on client outcomes. This includes quantitative and qualitative content analysis of therapist emails to clients receiving iCBT for depression [36], anxiety [69] or bulimia [87]. Here, Sanchez-Ortiz et al. [87] analysed 712 emails and found that 95.4% of all support emails included *supportive comments*, but little *cognitive or behavioral guidance* (<15%). Paxling et al. [69] studied 490 emails and found four support behaviors to positively correlate with module completion: *task reinforcement* —making positive references to what a client has already done or achieved in the program; *self-efficacy shaping* —prompting clients to engage in learned health promoting behaviors; *task prompting* —encouraging clients to complete the activities of the CBT program; and *empathetic utterances* —conveying an understanding of the person's suffering or life situation. Task reinforcement was further correlated with better client outcomes; whilst *deadline flexibility* (*e.g.* therapists postponing tasks) correlated negatively. Similar to task reinforcement, Holländare et al. [36] found (analysing 664 emails) *affirming* and *encouraging* behaviors (*e.g.*, validating and praising what the client did) most associated with immediate or longer-term improvements in outcomes, alongside therapist *self-disclosure*.

While previous research showed that the presence of a supporter correlates with better therapy outcomes, studies on the effectiveness of supporter behaviors remain sparse (*cf.* [36, 89]). Thus, there is a need for a deeper understanding of how supporter behaviors —as manifest in their online communications with clients —contribute to beneficial clinical outcomes.

Understanding Large-Scale Mental Health Behavior Data

Recent years have seen a growth in research exploring ML for mental health behavior data as captured by wearable health trackers [29]; mobile phones [10, 20, 22, 101, 104, 23]; social media [12, 44, 62, 68, 85]; or electronic healthcare records (EHR) [2, 98]. For the wealth of data that can be collected by these technologies, the fields of ML and data mining provide computational methods that can help improve our understanding of human behaviors and predicting or optimizing clinical outcomes [42]. Frequently applied methods that are particularly relevant to the approach taken in this paper are: clustering, text-mining and association rule mining (ARM).

Clustering & Text Mining in (Large-Scale) Mental Health Data

Clustering is an unsupervised ML technique to identify behavioral patterns in data through commonalities in each data piece; it is often used to identify features in unstructured data (*e.g.* [2, 14, 68]). In our research, two types of work are particularly relevant. Firstly, to better understand the behaviors of therapists engaged in text-based (SMS) counseling, Althoff et al. [3] clustered therapists, based on client outcomes, into groups of 'more' and 'less' successful counselors; and then compared how their communications differed using linguistic analysis. We followed a similar approach. Secondly, to identify support behaviors in thousands of anonymous supporter messages, we employ text-mining, which uses natural language processing (NLP) techniques to extract linguistic features or topics from large-scale text. In mental health, text-mining has been used to better understand discussion topics in online mental health communities [62, 68]; to study mental expressions [85], the receipt of social support [90], or cognitive change [74]. Few works seek to specifically aid moderators of online support communities in their work practices, *e.g.* by identifying helpful and unhelpful comments [12, 44]. Outside of social media, text mining is used to predict suicide risks from SMS messages [63], suicide notes [71], and EHRs [2] to aid care provision. Extending this body of work, we seek to better understand mental health support through a linguistic analysis of supporter messages as part of an iCBT intervention.

Using Association Rules for Behavioral Pattern Mining

Similar to clustering, association rule mining (ARM) is a common data mining technique for extracting behavioral patterns in data (*e.g.* [13, 23, 66, 104]). Here, the focus is on discovering interesting relations between variables in large data sets such as how patterns of certain data characteristics (*e.g.* client opinions, types of symptoms, demographics) relate to desirable outcomes (*e.g.* help-seeking behaviors, clinical score) [31, 66, 106]. We will show how we adapted an ARM algorithm to extract patterns of context-specific best practices of support.

THE I-CBT INTERVENTION

SilverCloud is an established iCBT platform for the treatment of depression, anxiety, and functional impairments. Its development builds on both HCI [21] and clinical research, including randomized controlled trials that evidence the clinical effectiveness of offered treatment programs [83]. In this paper, we focus on one of its most frequently used programs: treatment for depression and anxiety. Accessed online or via mobile, the program presents a self-guided intervention of seven core psycho-educational and psycho-therapeutic modules that are delivered using textual, video and audio contents as well as interactive activities, tools, quizzes and personal stories. Clients work through the program content at their own pace and time, with the recommendation to complete one module each week.

To encourage engagement and continued use, clients receive support from a trained supporter in the form of weekly reviews throughout their treatment journey. The supporters are graduate psychologists with further training in low-intensity interventions that are CBT based, including iCBT. Their support involves writing feedback messages to the client on their work, which usually takes 10-15 minutes to complete. Finally, to assess and monitor clients mental health throughout treatment, clients also complete clinical questionnaires each week, including the PHQ-9 for depression [47] and GAD-7 for anxiety [53]. Overall, the service aims to increase reach and provide effective care for those living with mental illness.

Frequency & Format of Supporter Interactions

Supporters typically review clients' work on a weekly basis over the 6-8 week treatment period. This serves to manage client expectations of support being restricted to certain times as opposed to immediate 24/7 availability. To this end, supporters can only see clients' activities and messages on, or after, the agreed review day. To review clients' progress, supporters have access to information about clients via usage metrics. These show: completed clinical scores of PHQ-9 and GAD-7; any client messages to the supporter; and how many content pages the client viewed, tools they used, and times they logged into the system. For each of these usage metric items, supporters can expand the view to retrieve more details about the content the client reviewed, and their responses in completing interactive tools. Clients have full transparency on what information is shared with their supporter [21].

In response to this client information, supporters compose a personalized feedback message. To facilitate this, they can select and adapt a messaging template from a drop-down menu within a specific supporter interface. These templates tend to be written by the supporters in their own words, and are then specifically tailored to each clients' situation. During training, supporters learn to personalize messages. Following prior research and guidelines this involves: referencing the clients name and things they did or said with a specific focus on activities of *task reinforcement* and *task prompting* [69]; *encouragement* [36, 82]; *guidance and advice* [82]; and effective communication using *simple language and explanations* [75]. As a final step, supporters can bookmark existing, and unlock additional therapy contents on the platform; and they select a subsequent review date. Once their message is submitted via the system, clients receive a notification email. They can view their supporter message when next accessing the program, at which time they will be automatically prompted to complete a new set of clinical questionnaires.

Dataset Description

Our dataset consists of information about: the supporter feedback messages; and the number and types of interactions that a client had with the platform (*e.g.*, how many CBT content pages they viewed); as well as the number and types of information that clients shared with their supporters (*e.g.*, number of journal entries, tool use data) in the time before and after each feedback message. Across the review period, we also have a subset of clinical scores indicative of the symptoms of the clients' depression (PHQ-9) and anxiety (GAD-7) before and after a supporter message. For our sample, initial PHQ-9 scores indicated that 32% of clients had minimal-to-mild symptoms of depression, 30% were moderate, 23% moderately-severe, and 15% severe. For GAD-7, initial scores showed that 36% of clients had minimal-to-mild symptoms of anxiety, 31% were moderate, and 33% severe. Typically, each client is assigned only 1 supporter, but if that supporter becomes unavailable, they may be assigned a new one. Table 1 contains basic dataset statistics.

Dataset statistics	
Supporters	3,481
Clients	54,104
Clients with >1 supporters	5,967
Messages	234,735
Messages with Pre & Post PHQ-9 & GAD-7	77,006
Average message length (in words)	191
Average message length (in sentences)	9.5

Table 1: Overview of basic dataset statistics.

To protect full anonymity of both clients and supporters, only non-person identifiable, high-level interaction data was used for the analysis. For clients this included numbers, types and frequencies of interaction events, and aggregates of clinical scores. For supporters this meant the number, frequency and length of each feedback message. Features extracted from message texts were restricted to coarse-grained linguistics to preserve anonymity. This matched the terms and conditions of the service, and user consent, which permits the analysis of anonymous data for research purposes, and to improve the effectiveness and service tools of the treatment platform.

Our research employs ML and data mining methods to better understand what support strategies (*e.g.* use of encouraging words) characterize supporter messages that are correlated with better clinical outcomes for clients. As a first step, this requires us to identify what constitutes 'successful support messages' based on clinical outcomes. To this end, we next describe: (i) how we defined clinical outcomes as change and improvement rates in clients over time; and (ii) then used these measures in clustering to achieve three clusters of supporters whose messages correlate with either 'high', 'medium', or 'low' success rates in improving clients' clinical scores; we use the 'high' and 'low' clusters in further analysis.

IDENTIFYING CLUSTERS OF SUCCESSFUL SUPPORT

To better understand how different support strategies in support messages correlate with better clinical outcomes for clients, we need to identify what constitutes an appropriate 'outcome measure' in relation to supporter messages. While clients are typically asked to complete PHQ-9 and GAD-7 questionnaires immediately following the receipt of a new supporter message, clients can complete the questionnaires at any time before their next scheduled supporter review. Further, for some supporter messages, an impact on client outcomes may not be apparent immediately (week-on-week); and instead requires a focus on how outcomes develop over time.

To measure clinical outcomes over a period of time, Randomized Controlled Trials (RCTs) tend to assess the difference between the clients' initial and final clinical scores [97, 41]. While feasible in study set-ups where clients complete an 'exit survey' after a fixed period of time, this is more complicated for real-world observational data that can include client drop-out. Further, client clinical outcomes can be highly dependent on their situation (*e.g.*, loss of a loved one), symptomatology (*e.g.* seasonal or cyclic depression), personality or learning style, meaning that some clients will not improve despite the use of good support strategies in supporter messages. However, good strategies used consistently with a set of clients, should result in improvement for the majority of clients in that set. Hence, in our analysis, we focus on the 'actors' (supporters) who employ the support strategies in their messages, instead of the 'receivers' (clients). To this end, we propose computing message-level and client-level clinical outcomes for each supporter in our dataset; we describe this next.

Change & Improvement Rates as Clinical Outcomes

We compute the following clinical outcomes by averaging post-message change in PHQ-9 and GAD-7 clinical scores across the messages sent by each supporter. Clients should complete these questionnaires in between messages. Messages with incomplete before or after questionnaires are excluded.

1. *Message-level Change (MC):* The clinical score after a message is highly dependent on the clinical score before the message, as clients that have more severe symptoms before the message also tend to improve more on an average after the message. Hence, we measure Message-level Change as the difference between actual change and the expected change given the client score before the message. That is, for each supporter S with NM messages, compute:

$$\frac{1}{NM}\sum_{r=1}^{NM}(actual\_change_m - expected\_change_m)$$
$$actual\_change_m = score\_before(m) - score\_after(m)$$
$$expected\_change_m = score\_before(m) - \mathbb{E}(score\_after(m)|score\_before(m))$$

2. *Message-level Improvement Rate (MR):* If Message-level Change > 0, then the client improved more than expected post-message, and we label the message as "improved (1)". Otherwise, we label the message as "not improved (0)". For each supporter S with NM messages, we average these labels across all messages to compute this outcome.

3. *Client-level Change (CC):* While MC captures changes in clinical scores across all messages by a supporter, CC normalizes these changes across all clients of the supporter. For each supporter S, we first compute the MC for each client of S separately using the messages that S sent to them. Then, we average the MCs per client across all clients of S to get CC. *E.g.*, if a supporter sends 6 messages to client A whose change is $+1$ after each messages and 4

messages to client B whose change is always 0, the MC will be $\frac{6}{6+4} = 0.6$ while the CC will be $\frac{(6/6)+(0/4)}{2} = 0.5$. Thus, MC can be high even when only a few clients improve, whereas CC will only be high when these improvements are consistent across all/ many clients. This makes CC more robust to a single client's changing situations or symptoms.

4. *Client-level Improvement Rate (CR):* For each supporter S with clients NC, we first compute the average Message-level Improvement Rate using messages S sent to each client separately, and then sum these rates across all clients and divide by the total number of clients.

Clustering Supporters Based on Support Outcomes

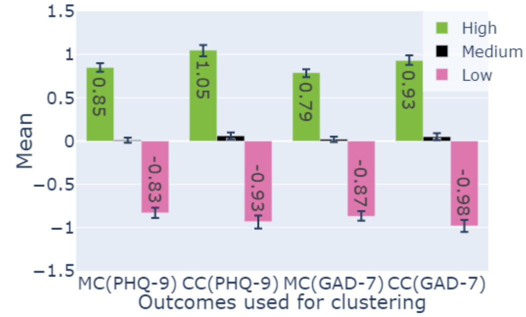
When computing the above 4 outcome measures for both PHQ-9 and GAD-7, we achieve 8 outcome measures per supporter. As a next step, we use these 8 measures as 'features' for clustering. We apply K-means clustering to obtain K=3 clusters of supporters whose messages are generally linked with either 'high', 'medium' or 'low' improvements in client outcomes. The number of k=3 clusters was determined by plotting the sum of squared distances of samples to their closest cluster center against the number of clusters, and visually inspecting the elbow obtained. *We hypothesize that there are differences in the support messages sent by supporters in the 'high' versus 'low' outcome clusters; and that these differences will help us identify what may constitute more effective support strategies*. Table 2 reports the summary statistics of the three obtained clusters of supporters. Figure 1 shows the mean values of all 8 outcomes measures in these clusters, along with the 95% bootstrapped confidence intervals as error bars. Given the mean values of the outcomes and the narrow 95% confidence intervals, we can see that our clustering has reliably divided the supporters into 3 robust groups where these outcomes are *typically*: high, medium, and low. We did additional statistical analysis that confirmed the results in Figure 1 and showed how, independent of clients initial clinical scores, the differences in mean PhQ-9 and GAD-7 scores between the high and low clusters was significant ($p<0.05$).

Cluster	#Supporters	#Clients	#Messages	#Messages Labeled
High	438	11068	42734	14519
Medium	767	31789	123303	42740
Low	393	10828	47023	14266

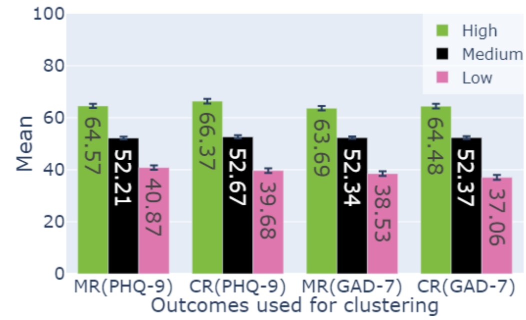
Table 2: Statistics for the 'high', 'medium' and 'low' outcome clusters. Only supporters with >9 (Median value) labeled messages for PHQ-9 and GAD-7 were included in the clustering.

IDENTIFYING SUCCESSFUL SUPPORT STRATEGIES

As a next step, we want to identify what semantic or linguistic strategies are frequently associated with 'high' client outcomes. For this purpose, we analyze the differences between the messages sent by supporters in the 'high' outcomes cluster and those in the 'low' outcomes cluster. We are interested in identifying support strategies that occur significantly more often in the messages of supporters in the 'high' outcome cluster. Further, this difference needs to be consistent across different client contexts; meaning that the result has to be independent



(a) Message-level (MC) and Client-level (CC) Changes



(b) Message-level (MR) and Client-level (CR) Improvement Rates

Figure 1: Means of all 8 outcomes in each supporter cluster with the 95% bootstrapped confidence intervals as error bars.

from variations in individual client context variables such as: the extent to which a client engages with the iCBT program ($ContentViews$), shares content with their supporter ($Shared$), the sequence number of the message *i.e.* the number of messages received so far plus one: $MessageNumber$), or clients' current mental health state ($CurrentPHQ-9$ and $GAD-7$).

Methodology: <Strategy> Across <Context> Bins

We are mindful that the actions of supporters, as manifested in their feedback messages to clients, present a direct response to what they know about their clients' situation (*e.g.* symptom severity, level of platform use). To disentangle the clients' context from, whilst understanding the role that context can play in the use of, specific support strategies, we decided to first divide the messages of supporters in the 'high' and 'low' outcome clusters into different 'data bins' for each of 5 client context variables. This allows us to compare the differences in strategies found in support messages of the 'high' and 'low' clusters separately for different client contexts; and thereby to assess, if identified significant differences between the two groups are consistent across, and independent of, variations in client contexts. We found this approach to be more feasible than the use of context as a control variable in linear regression, which due to the large size of the data sample frequently resulted in statistically significant results, but whose effect sizes were difficult to interpret.

Next, we detail on: (i) the client *context* variables that we defined; (ii) semantic or linguistic features we extracted from supporter messages as *strategy* variables; and (iii) how we analyzed what semantic or linguistic support strategies were salient in the 'high' versus 'low' support clusters across various bins of each client *context* variable.

Client <Context> Variables

For each message i of every client, we extracted 5 *context* variables that describe the client's situation *just before the message i.e. between the* $(i-1)^{th}$ *and* i^{th} *messages*:

(1) *ContentViews*: Number of content views *i.e.* number of times the client clicked on a module or topic in the iCBT program (10 Bins[1]: 0 views, 8 bins of 1-80 views in increments of 10 views, and >=81 views).

(2) *Shared*: Number of data instances shared with the supporter *i.e.* the total number of tools and journal entries shared, and messages sent to the supporter (4 Bins: 0 shared instances, 3 bins of 1-15 in increments of 5. We never had >=16.).

(3) *MessageNumber*: clients are expected to received up to 8 reviews. For the message number i we have 9 bins (8 bins for $i = 1$ to 8, and 1 bin for $i >= 9$).

(4-5) $CurrentPHQ-9$ and $CurrentGAD-7$: Clinical scores measuring the client's depressive and anxious symptoms, as seen by the supporter before the i^{th} message. Higher scores indicate more severe symptoms (7 Bins[1]: score $<= 9$, 5 bins for scores 10-19 in increments of 2, and score $>= 20$).

Semantic or Linguistic <Strategy> Variables in Messages

Similar to text-mining approaches for online mental health content analysis [12, 44, 62, 68, 85], a lexicon-based approach was used to extract comprehensive, high-level language characteristics from the supporter messages without risking to identify any text content. For each supporter message i, NLP techniques were used to extract 23 features indicative of a support strategy (*e.g.*, use of encouraging phrases). The features were achieved by mapping each word in the message i to a word category of a lexicon. We defined 23 *strategy* variables based on the literature and available lexicons:

(1-3) Sentiment: To capture the overall *sentiment* of a message, we used the NRC Emotion Lexicon [57] to extract the percentages of positive (*Pos*) and negative words (*Neg*), and the difference between them ($Pos-Neg$). Word percentage were used instead of absolute word counts, to better compare messages of different lengths [92].

(4-11) Emotions: As an indicator of the *emotional tone* of a message, we extracted the percentages of words related to the 8 emotion categories of the same lexicon [57]: *Anger*, *Anticipation*, *Disgust*, *Fear*, *Joy*, *Sadness*, *Surprise*, *Trust*.

(12-13) Pronouns: We extracted the percentage of first person plural pronouns (*e.g. We*, us, our), assuming that supporters will use these to convey a supportive alliance [9]. Assuming second person pronouns (*e.g.* you) are used more often to direct clients to engage in specific program activities (e.g., task prompting behaviors [69]), we calculated the difference between the percentages of first person plural and second person pronouns ($We-You$) as indicator of *supportive alliance*.

[1]Increments of bins were heuristically chosen by rounding up half the standard deviation, after excluding the first and last bins.

(14) Encouraging Phrases: Based on a series of support message examples used in supporter training, we derived a list of 15 commonly used encouraging phrases (*e.g.* 'good job', 'well done'). As an indicator of the *motivational tone* of a message, we calculated the ratio of the number of encouraging phrases to the number of sentences overall ($Encouragement : Sentences$).

(15-22) Mental Processes and Behaviors: We used the Regressive Imagery Dictionary [54], an extensively used lexicon in mental health research (*e.g.* to analyze psychotherapy dialogue [80, 37, 93]), for analyzing different categories of conceptual thought in text. Specifically, the analysis includes the seven categories of secondary cognition that are *logical, reality-based, and focused on problem solving*. This includes the percentages of words related to mental processes of: *Abstraction* (*e.g.*, know, may, thought), *InstrumentalBehavior* (*e.g.*, make, find, work), *MoralImperative* (*e.g.*, should, right, virtue), *Order* (*e.g.*, simple, measure, array), *Restraint* (*e.g.* must, stop, bind), *SocialBehavior* (*e.g.*, say, tell, call), and *TemporalReferences* (*e.g.*, when, now, then). As additional behavioral cues, the percentage of *ActionVerbs* in the text were also extracted, as specified by the Moby Project [73].

(23) Quantity: As a measure of *quantity of support*, we assess the length of the text messages via number of words (*WordCount*) and report number of sentences where relevant.

Analysis of Strategies Across Contexts

We identify a semantic or linguistic support strategy as successful, if: (i) it occurs significantly more often in messages by supporters in the 'high' outcome cluster (compared to the 'low' outcome cluster); and (ii) this difference is consistent across different client contexts. We operationalize this definition by considering each $< context, strategy >$ pair separately in our analysis. Since we have extracted 5 *context* variables and 23 *strategy* variables, we analyze 115 $< context, strategy >$ pairs. For each pair, we do the following:

1. We divide the messages from the 'high' and 'low' supporter clusters into multiple bins according to the client's *context* before the supporter composed a message.
2. For each bin, we then compute the mean of the *strategy* variable for composed messages in the two clusters *separately* along with their 95% bootstrapped confidence intervals.
3. For each bin, we also compare the means of the *strategy* variable across the two clusters using a bootstrapped resampling test.
4. Relevant support strategies should have statistically significant differences in means between the two clusters (α = 0.05) and 95% confidence intervals that rarely overlap, across most bins.

Since the messages in each bin can belong to the same client or supporter, they are not independent. To address this, we use hierarchical bootstrapping for clustered data by randomly sampling supporters with replacement, and their clients and reviews without replacement 10000 times [78]. For comparing the means of the messages in each bin across the two clusters, we reject H_0: Difference in means = 0 (two-tailed) at level

α if and only if a valid confidence interval with coverage probability $1 - \alpha$ does not cover 0 (α=0.05 in our analysis).

While only messages with pre and post PHQ-9 and GAD-7 scores were used in the initial clustering, we used all messages by supporters in the 'high' and 'low' clusters for this analysis, so as to not miss out on clients with low levels of usage.

Results: Strategies Used in Successful Messages

What follows are the main findings of significant differences in the linguistic strategies that were used in supporter messages associated with 'high' versus 'low' outcomes (improvements in client clinical scores) across different client contexts.

Sentiment and Emotion in Support Messages

Across *all* supporter messages in our dataset, the sentiment analysis showed that positive words were generally used more frequently than negative words ($Mean_{pos}$ = 6.2%, SD_{pos} = 2.8; $Mean_{neg}$ = 1.7%, SD_{neg} = 3.4; $Mean_{pos\text{-}neg}$ = 4.5%, $SD_{pos\text{-}neg}$ = 1.3). Further, **more successful supporter messages consistently used more positive and less negative words**. This effect remained consistent (*i.e.* $p < 0.05$ for all 3 sentiment *strategy* variables across all bins of the 5 *context* variables).

We also found that **more successful messages had less occurrences of negative emotions conveying sadness and fear than less successful messages** (*i.e.* $p < 0.05$ for these two emotion-related *strategy* variables across all bins of the 5 *context* variables). In addition, more successful messages used more frequently words that expressed joy, yet the difference in means was non-significant for several bins. Since anger and disgust rarely featured in the messages, we did not include them in further analysis, and there were no statistically significant results for any of the other emotions.

Pronouns: Sense of Supportive Alliance

Our results show that, across all messages, second person pronouns (*e.g.* you) were used more frequently than first person plural pronouns (*e.g.* we, us, our). However, we found that **more successful messages consistently employed first person plural pronouns more frequently than less successful messages, and had greater differences between uses of first person plural and second person pronouns** (*i.e.*, $p < 0.05$ for both pronoun *strategy* variables across all bins of the 5 *context* variables). Consistent with previous work, first person plural pronoun use may reflect a sense of supportive affiliation [30] or increased presence in computer-mediated communication [46].

Encouraging Phrases: Motivational Tone

To assess the motivational tone of a message, we calculated the ratio of number of common encouraging phrases to sentences (*Encouragement* : *Sentences*). Across all dataset messages, the mean ratio is 0.04 (*SD* = 0.09). Given the average length of a support message of 9.5 sentences, we can estimate that an encouraging phrase is used, on average, once in every 2.6 messages. We found that **more successful messages consistently contained significantly more encouraging phrases compared to less successful messages** ($p < 0.05$ for this *strategy* variable across all bins of the 5 *context* variables except the 61-70 bin for *ContentViews*).

Mental Processes & Behaviors: Action Orientation

For our our mental process variables, we found that **more successful messages consistently employed more words associated with social behavior and less words associated with abstraction, when compared to less successful messages** ($p < 0.05$ for these *strategy* variables across all bins of the 5 *context* variables). In order to better interpret the results of our analysis, we visualize each of these $< context, strategy >$ pairs. Figure 2 (left) shows the percentage of words associated with *SocialBehavior* that are used in 'more' and 'less' successful support messages. More specifically it plots the mean of this *strategy* variable for all messages in each bin for the 'high' and 'low' outcome clusters on the Y-axis, and for each bins of the *ContentViews* context variable on the X-axis. The error bars are 95% bootstrapped confidence intervals of the means. Some of the most frequently used social behavior words were: *help, call, discuss*, and *share*.

Figure 2 (right) further shows how 'more' successful messages contained on average less words associated with the strategy variable *Abstraction* than 'less' successful messages; and this finding was consistent independent, *e.g.*, of the clients' depressive symptoms prior to the supporter review (*cf.* the bins of the $CurrentPHQ-9$ variable). Frequently used abstraction words were: *think/thought, know, understand*, and *learn*.

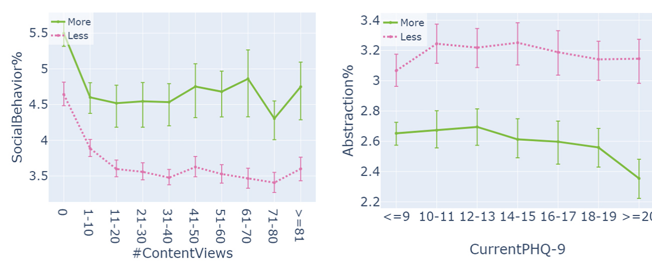


Figure 2: Mean percentage of words in 'more' and 'less' successful support messages that are associated with: *SocialBehavior* across each bin of the *ContentViews* context variable (left); and *Abstraction* across each bin of the $CurrentPHQ-9$ context variable (right).

Restraint, *MoralImperative*, and *Order* occurred rarely in our dataset (as indicated by their mean occurrence percentages), and thus were not included for further analysis. For words associated with *InstrumentalBehavior*, the results were not statistically significant. While the results for words associated with *TemporalReferences* were statistically significant across most bins of the *context* variables, the 95% confidence intervals overlapped and did not show any consistent trend.

Lastly, we found that, on average, about a quarter of every message comprised of action verbs (*Mean* = 24.4%, *SD* = 4.5), and that **more successful messages consistently employed more action verbs compared to less successful ones** (*i.e.* $p < 0.05$ across all bins of the 5 *context* variables).

Quantity: Length of Supporter Messages

Finally, we found that messages were on average 9.5 sentences long (*SD* = 6.7) and included 191 words (*SD* = 147). We found that **more successful messages were consistently**

shorter ($Mean_{words}$ = 177, $Mean_{Sentences}$ = 9) **than less successful ones** ($Mean_{Words}$ = 214, $Mean_{Sentences}$ = 10.4); with $p < 0.05$ for *Word Count* across all bins of the 5 *context* variables with very few exceptions). This surprisingly contradicts previous findings by [3], and thus, requires further research.

CONTEXT-SPECIFIC PATTERNS OF SUPPORT SUCCESS

In the previous section, we identified support strategies that were consistently associated with better clinical outcomes across 5 types of contexts, which we each had treated in isolation. In this final analysis, we want to better understand the more complex relationship that likely exists between the use of support strategies and different client context variables. In other words, how may considering the combination of multiple context variables (rather than each client *context* variable by itself) shift how salient a specific support strategy is in messages associated with either 'high' or 'low' client outcomes. We believe that identifying such relationship patterns could enable a more effective tailoring of support strategies to specific client contexts. Next, we therefore describe our approach to identifying <multidimensional context → strategy> patterns.

Extracting Multidimensional Context → Strategy Patterns

To identify <multidimensional context → strategy> patterns, we needed to discover associations between client contexts containing two or more contextual variables, and each support strategy. To find associations among multiple items in datasets, we used the well-known frequent item set mining *Apriori* algorithm. It first generates a set of frequent items that occur together, and then extracts association rules that explain the relationship between those items. Given two sets of items A and B, an association rule $A \rightarrow B$ exists if A and B frequently appear together in transactions. In our case, "transactions" are supporter messages that occur in client context A and employ strategy B. We consider each strategy separately in various client contexts, such that context A is multidimensional (it contains two or more of the 5 *context* variables) while strategy B will always be one singular *strategy* variable[2].

We extracted association rules separately for both our 'high' and 'low' outcome clusters; and then compared the rules that occurred in both groups using two "interestingness" measures: Confidence and Coverage [33, 49]. Confidence is the percentage of transactions containing A that also contain B ($P(B|A)$), whereas coverage is the percentage of transactions containing A ($P(A)$), chosen as a measure of the applicability of a rule. For example, if a context A occurs very frequently (*e.g.* coverage is high), rules associated with it are more interesting as they can be applied more often when recommending support strategies personalized to the client's context. We set minimum coverage to 0.01 and minimum confidence to 0 to extract a large number of rules, that we can then filter using a different metric that compares the salience of the rules.

For this analysis, we excluded the strategy variables: *Anger*, *Disgust*, *Restraint*, *Order*, and *MoralImperative*, as they rarely occurred in our data. The remaining 18 *strategy* variables were discretized to 'high', 'medium', and 'low', using equal frequency binning[3]. Then, Apriori was applied separately on messages from our 'high' and 'low' outcome clusters. We found 22599 rules in both clusters. The salience for each of these rules was calculated as the absolute confidence difference between the two clusters. That is, **the salience of a rule $A \rightarrow B$ is measured as the difference between percentage of strategy B occurring across the two clusters, when the client is in context A**[4], such that more salient rules are used more frequently by supporters in either the more or less successful cluster. We select a subset of rules to interpret by first choosing interesting contexts and strategies that are used in at least one rule with salience >= 0.20, and then selecting all rules associated with them (giving us 1584 rules associated with 66 contexts and 8 *strategy* variables with 3 levels each[5]), allowing us us to interpret the most salient rules in context.

Results: Salient Context-Specific Support Strategies

Figure 3 shows a heat map of the salience of these rules. Next, we present a few examples of interesting results. Taking the third row from the bottom, read horizontally, we see the rules associated with low *WordCount*. The row is mostly green, apart from the first 6 context rules. The green indicates that the support strategy "*WordCount* = low" is more salient in more successful support messages, with a darker shade indicating higher salience. We see that for almost all contexts, a lower word count is more salient in more successful messages. However, this is flipped for the first 6 context rules that show no client engagement prior to a review (*e.g. SharedContent* with supporter = 0, *ContentViews* = 0). Thus, for disengaged clients, writing shorter messages is more strongly associated with less successful outcomes (shown as color pink).

Outside of 'salience flips' between the two groups (color change); interesting patterns can also be identified through variations in the shade of the same color. For example, for strategies "*Fear* = low" and "*We* = high", we see high salience (dark green) in more successful messages for the same first 6 client contexts; and reduced effects thereafter. This means that writing messages with less words related to fear, and more first person plural pronouns are stronger associated with more successful support messages, and this effect is particularly salient in situations where clients are disengaged.

Aligned with our previous results, we further find for highly engaged clients that successful messages reference more social behavior (*SocialBehavior%*=high) and less abstraction (*Abstraction%*=low), while less successful messages reference less social behavior (*SocialBehavior%*=low or medium) and more abstraction (*Abstraction%*=high) in the same contexts. There remain more rules and patterns to unpack. Here, we see our analysis as contributing initial insights on how identified support strategies can be more or less salient or successful depending on a specific client context. So far, our results imply that **for less engaged clients, writing longer, more positive and more supportive reviews is linked with**

[2] The arules R package was used to set up Apriori in this way [38].

[3] Messages from the 'high', 'low', and 'average' clusters were used for discretization.

[4] Salience = abs($P(B_{high}|A_{high})$-$P(B_{low}|A_{low})$)

[5] Rules containing *ActionVerb%* were excluded, because action verbs occur too frequently and are not very informative

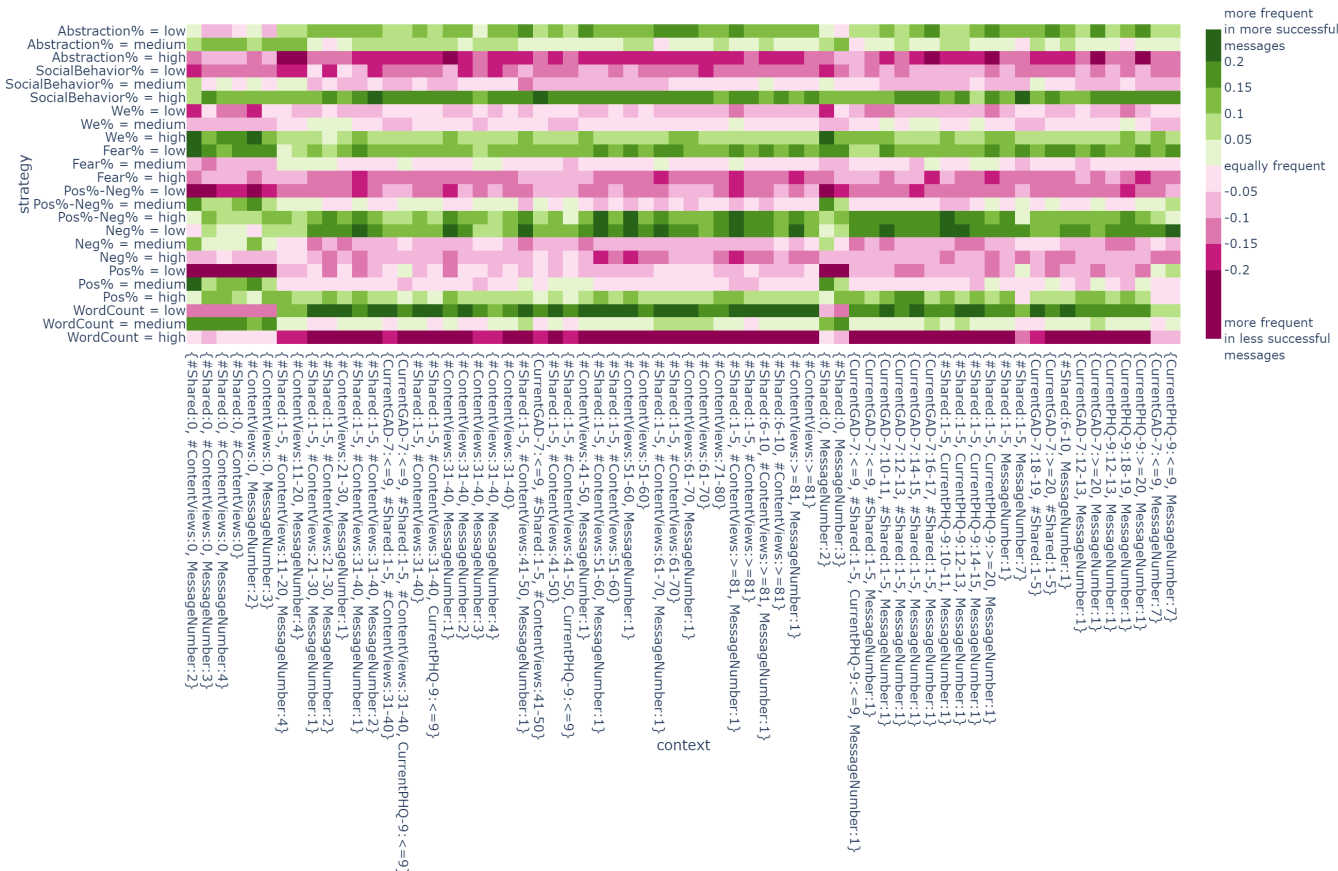


Figure 3: The salience of the most interesting <multidimensional context (A) $\rightarrow$ strategy (B)> patterns are visualized. Each cell represents the salience of a pattern $A \rightarrow B$. Context A is on the X-axis while strategy B is on the Y-axis. Salience is the difference between percentages of rules containing B given A across the two clusters, and is multiplied by "-1" when the rule occurs more frequently in less successful messages. Thus, *rules in green occur more frequently in more successful messages while rules in pink occur more frequently in less successful messages. Darker colors imply greater salience.* For readability, the contexts are sorted on *ContentViews* followed by *Shared* and then *MessageNumber*. The figure is best read *strategy*-wise from left to right.

greater outcomes; whilst **more engaged clients appear to benefit more from messages with less negative words, less abstraction, and more references to social behaviors**.

DISCUSSION

Our work presents the first application of unsupervised ML, and statistical and data mining methods to analyze complex, large-scale supporter-client interaction data of an established iCBT intervention for the treatment of depression and anxiety. We focused on developing a better understanding of how the behaviors of supporters, who assist clients' engagement with this service, may correlate with better clinical outcomes for these clients; which presents a largely under-explored area. Below, we discuss the main implications of our work for future research that intersects the fields of HCI, ML and healthcare.

Identifying Effective Context-Specific Support Strategies

We described our approach to identifying 'more' and 'less' successful support behaviors that are manifest in communications to clients. Using semantic and linguistic feature extraction methods, our results indicate that supporter messages that typically achieve higher client outcomes contain more words that are positive, supportive, related to social behaviors, and less abstract; and those messages tend to be shorter than less successful messages. Largely, these findings align well with previous qualitative studies of iCBT support that emphasize the prevalence of supportive language [87], and importance of affirmations and encouragement [36, 69] for client outcomes. Extending this research in iCBT, our work presents novel findings of how the success of identified support strategies can vary dependent on a specific client context. Next, we discuss how having a better understanding of each persons' context for support enables new opportunities for personalized care; which, in turn, can improve client engagement with iCBT interventions and benefit their mental health outcomes.

Data-Enabled, Personalized Mental Health Support

So far, only few works have explored the design space and use of ML to personalize the treatment or delivery of (mental) health interventions (*e.g.* [24, 67, 60]). Most prominently, Paredes et al. [67] applied reinforcement learning for recommending stress coping interventions tailored to the person. Other recent trends include the development of just-in-time adaptive interventions (JITAIs) that utilize algorithms to provide the right type or amount of support, at the right time, in response to an individuals' internal and contextual state [60].

Design Implications for Personalized Human Support in iCBT

For guided iCBT interventions, there are multiple ways in which gathered data insights about context-specific support strategies can inform supporter interface design. For example, as supporters review a client, they may be presented in their feedback interface with recommendations of what strategies specific to this clients' situation may be strongly correlated with successful client outcomes, providing them additional input to their feedback practices. To help translate linguistically-derived support strategies more directly into feedback messages, strategy-related words (*e.g.* positive words, certain pronouns) could be highlighted in real-time as part of the message editor that supporters use. Especially for training purposes and to support skills acquisition of novice supporters, it may also be helpful to integrate examples of 'support messages that were successful for similar clients' as guidance.

Human-Centered ML in (Mental) Healthcare

The above design examples further illustrate our orientation to human-centred ML and the integration of data insights into healthcare practice. Rather than promoting the use of templates or automatizing client feedback away from the human supporter, we suggest designing interventions that seek to *enhance supporter agency* by enabling them to personalize their feedback more effectively for each person (*cf.* [94]), and to better understand how their actions make (ideally) a positive difference to their clients. Thus, while advanced data tools that can identify complex patterns are often seen to generate more accurate, objective and less biased insights (*cf.* [34, 98, 105]), it is important that we (i) do not take away from, but help foster, the important relationship and genuine *human connection* that is formed between supporter and client and crucial to their alliance and positive outcomes [9, 58]; and (ii) ensure supporters feel that their input and expertise is valued rather than made redundant or replaced in favor of data science.

Understanding (Big) Data in Digital Health Interventions

Next, we discuss identified challenges and opportunities for working with complex, large-scale observational data.

Trade-Offs between Data Access and Use & Ethical Conduct

Although we had unprecedented, privileged access to large-scale supporter-interaction data, we made necessary trade-offs as to what kind of analysis we could conduct to protect the full anonymity of both client and supporter data. This meant much of our analysis was restricted to coarse-grained linguistic features and high-level usage data. While our research captured individual word associations, the use of other linguistic features such as n-grams (*e.g.* [2, 44]) could expand the analysis to word pairings, or even short word sequences that could enable a richer contextual understanding of identified support behaviors. At the same time, however, such explorations need to be done with care so as to not risk making too much content, or the people who produced it, identifiable.

Caution is also required for secondary data analysis for which additional user consent is likely unfeasible to collect (for every subsequent analysis). As is common in ML approaches for mental health, user privacy should be carefully addressed to preserve anonymity (cf. phone data [10, 101], sensors [77], social media [68, 85, 105]), and analysis should occur in a context where there is a clear case for the prospective benefits that could arise, *e.g.*, from improved healthcare provision (see public health ethics [15] and recent work on user acceptance of digital phenotyping [52, 84]). Going forward, we need to continue developing feasible, privacy-preserving data methods; and, as researchers, need to remain critical of, and sensitive to, the extent to which our data analysis is really 'justifiable' with regards to how it comes to benefit users and health services.

Deriving Insight from Data: Interpretation & Future Directions

Due to the necessary data use restrictions, we acknowledge that derived data insights —whilst novel —are limited to the definitions chosen, and require further research to validate. Future work may also explore multiple additional data mining avenues, including: (i) analysis of supporters' use and adaptation of messaging templates (*cf.* [3]); (ii) studies into sequential routines of strategies (*e.g.* using Inverse Reinforcement Learning [7]); (iii) supporter clustering using 'engagement' as an outcome metric alongside clinical improvement (*cf.* [59]); or (iv) the combination of support strategies, supporter-features, and client engagement features to predict clinical outcomes.

Despite manifold possibilities for data mining, challenges remain for ensuring that derived insights are both *human interpretable* (*e.g.* [1, 35, 72]) and *clinically useful*. While we were deliberate in our choice of data tools and visualizing of their results to create representations that are comprehensive to laypeople, other research methods (*e.g.* qualitative studies [36, 69, 87]) are needed to help contextualize, validate and advance our understanding of support, or other data-derived health behaviors. More research is also needed to develop our understanding of the potential value that these new types of data insights could bring to actual healthcare practices.

CONCLUSION

Aiming to understand how the behaviors of iCBT supporters impact client outcomes, we presented our ML approach and the analysis of 234,735 supporter messages sent to an unprecedentedly large clinical sample of 54,104 clients with a mental health diagnosis. Using various computational methods we identified support behaviors associated with 'more' or 'less' improvements in clinical scores, and showed how their salience varied dependent on different client contexts. Our work enables a better understanding of best practices, and opens-up new opportunities for personalizing support provision within computer-delivered mental health services. We discussed: the implications of our findings for the design of iCBT supporter interfaces; the need for a human-centered approach for sensibly integrating data insights into healthcare practices; and ethical considerations for secondary data use.

ACKNOWLEDGEMENTS

We thank our broader team: James Bligh, David Carter, Dessie Keegan, Aditya Nori, Tim Regan, Sebastian Tschiatschek.

REFERENCES

[1] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y Lim, and Mohan Kankanhalli. 2018. Trends and trajectories for explainable, accountable and intelligible systems: An hci research agenda. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. ACM, 582.

[2] Marios Adamou, Grigoris Antoniou, Elissavet Greasidou, Vincenzo Lagani, Paulos Charonyktakis, and Ioannis Tsamardinos. 2018. Mining free-text medical notes for suicide risk assessment. In *Proceedings of the 10th hellenic conference on artificial intelligence*. ACM, 47.

[3] Tim Althoff, Kevin Clark, and Jure Leskovec. 2016. Large-scale analysis of counseling conversations: An application of natural language processing to mental health. *Transactions of the Association for Computational Linguistics* 4 (2016), 463–476.

[4] Gerhard Andersson, Nickolai Titov, Blake F Dear, Alexander Rozental, and Per Carlbring. 2019. Internet-delivered psychological treatments: from innovation to implementation. *World Psychiatry* 18, 1 (2019), 20–28.

[5] G Andrews, A Basu, P Cuijpers, MG Craske, Peter McEvoy, CL English, and JM Newby. 2018. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. *Journal of anxiety disorders* 55 (2018), 70–78.

[6] Nina B Baltierra, Kathryn E Muessig, Emily C Pike, Sara LeGrand, Sheana S Bull, and Lisa B Hightow-Weidman. 2016. More than just tracking time: complex measures of user engagement with an internet-based health promotion intervention. *Journal of biomedical informatics* 59 (2016), 299–307.

[7] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff, and Anind K Dey. 2016. Modeling and understanding human routine behavior. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 248–260.

[8] Jakob E Bardram, Mads Frost, Károly Szántó, Maria Faurholt-Jepsen, Maj Vinberg, and Lars Vedel Kessing. 2013. 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*. ACM, 2627–2636.

[9] Edward S Bordin. 1979. The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, research & practice* 16, 3 (1979), 252.

[10] Bokai Cao, Lei Zheng, Chenwei Zhang, Philip S Yu, Andrea Piscitello, John Zulueta, Olu Ajilore, Kelly Ryan, and Alex D Leow. 2017. Deepmood: modeling mobile phone typing dynamics for mood detection. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 747–755.

[11] Alan Carr. 2008. *What works with children, adolescents, and adults?: a review of research on the effectiveness of psychotherapy*. Routledge.

[12] Stevie Chancellor. 2018. Computational Methods to Understand Deviant Mental Wellness Communities. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, DC05.

[13] Rosa Chaves, Javier Ramírez, JM Górriz, Carlos García Puntonet, Alzheimer's Disease Neuroimaging Initiative, and others. 2012. Association rule-based feature selection method for Alzheimer's disease diagnosis. *Expert Systems with Applications* 39, 14 (2012), 11766–11774.

[14] Annie T Chen, Shuyang Wu, Kathryn N Tomasino, Emily G Lattie, and David C Mohr. 2019. A multi-faceted approach to characterizing user behavior and experience in a digital mental health intervention. *Journal of biomedical informatics* 94 (2019), 103187.

[15] James F Childress, Ruth R Faden, Ruth D Gaare, Lawrence O Gostin, Jeffrey Kahn, Richard J Bonnie, Nancy E Kass, Anna C Mastroianni, Jonathan D Moreno, and Phillip Nieburg. 2002. Public health ethics: mapping the terrain. *The Journal of Law, Medicine & Ethics* 30, 2 (2002), 170–178.

[16] Jonathan E Cook and Carol Doyle. 2002. Working alliance in online therapy as compared to face-to-face therapy: Preliminary results. *CyberPsychology & Behavior* 5, 2 (2002), 95–105.

[17] David Coyle and Gavin Doherty. 2009. Clinical evaluations and collaborative design: developing new technologies for mental healthcare interventions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2051–2060.

[18] David Coyle, Gavin Doherty, Mark Matthews, and John Sharry. 2007. Computers in Talk-based Mental Health Interventions. *Interact. Comput.* 19, 4 (July 2007), 545–562. DOI: http://dx.doi.org/10.1016/j.intcom.2007.02.001

[19] Marco de Sá and Luís Carriço. 2012. Fear therapy for children: a mobile approach. In *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 237–246.

[20] Orianna DeMasi and Benjamin Recht. 2017. A step towards quantifying when an algorithm can and cannot predict an individual's wellbeing. 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*. ACM, 763–771.

[21] Gavin Doherty, David Coyle, and John Sharry. 2012. 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*. ACM, 1421–1430.

[22] Afsaneh Doryab, Mads Frost, Maria Faurholt-Jepsen, Lars V Kessing, and Jakob E Bardram. 2015. Impact factor analysis: combining prediction with parameter ranking to reveal the impact of behavior on health outcome. *Personal and Ubiquitous Computing* 19, 2 (2015), 355–365.

[23] Afsaneh Doryab, Daniella K Villalba, Prerna Chikersal, Janine M Dutcher, Michael Tumminia, Xinwen Liu, Sheldon Cohen, Kasey Creswell, Jennifer Mankoff, John D Creswell, and others. 2019. Identifying Behavioral Phenotypes of Loneliness and Social Isolation with Passive Sensing: Statistical Analysis, Data Mining and Machine Learning of Smartphone and Fitbit Data. *JMIR mHealth and uHealth* 7, 7 (2019), e13209.

[24] Malin Eiband, Hanna Schneider, Mark Bilandzic, Julian Fazekas-Con, Mareike Haug, and Heinrich Hussmann. 2018. Bringing transparency design into practice. In *23rd International Conference on Intelligent User Interfaces*. ACM, 211–223.

[25] Sindhu Kiranmai Ernala, Michael L Birnbaum, Kristin A Candan, Asra F Rizvi, William A Sterling, John M Kane, and Munmun De Choudhury. 2019. Methodological gaps in predicting mental health states from social media: Triangulating diagnostic signals. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 134.

[26] Gunther Eysenbach. 2005. The law of attrition. *Journal of medical Internet research* 7, 1 (2005), e11.

[27] Kathleen Kara Fitzpatrick, Alison Darcy, and Molly Vierhile. 2017. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR mental health* 4, 2 (2017), e19.

[28] Marzyeh Ghassemi, Tristan Naumann, Peter Schulam, Andrew L Beam, and Rajesh Ranganath. 2018. Opportunities in machine learning for healthcare. *arXiv preprint arXiv:1806.00388* (2018).

[29] Martin Gjoreski, Hristijan Gjoreski, Mitja Luštrek, and Matjaž Gams. 2016. Continuous stress detection using a wrist device: in laboratory and real life. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 1185–1193.

[30] Erving Goffman. 1981. *Forms of talk*. University of Pennsylvania Press.

[31] Dion H Goh and Rebecca P Ang. 2007. An introduction to association rule mining: An application in counseling and help-seeking behavior of adolescents. *Behavior Research Methods* 39, 2 (2007), 259–266.

[32] Heather D Hadjistavropoulos, Nicole E Pugh, Hugo Hesser, and Gerhard Andersson. 2017. Therapeutic alliance in internet-delivered cognitive behaviour therapy for depression or generalized anxiety. *Clinical psychology & psychotherapy* 24, 2 (2017), 451–461.

[33] Jiawei Han, Jian Pei, and Micheline Kamber. 2011. *Data mining: concepts and techniques*. Elsevier.

[34] Tad Hirsch, Christina Soma, Kritzia Merced, Patty Kuo, Aaron Dembe, Derek D Caperton, David C Atkins, and Zac E Imel. 2018. It's hard to argue with a computer: Investigating Psychotherapists' Attitudes towards Automated Evaluation. In *Proceedings of the 2018 Designing Interactive Systems Conference*. ACM, 559–571.

[35] Fred Hohman, Andrew Head, Rich Caruana, Robert DeLine, and Steven M Drucker. 2019. Gamut: A design probe to understand how data scientists understand machine learning models. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 579.

[36] Fredrik Holländare, Sanna Aila Gustafsson, Maria Berglind, Frida Grape, Per Carlbring, Gerhard Andersson, Heather Hadjistavropoulos, and Maria Tillfors. 2016. Therapist behaviours in internet-based cognitive behaviour therapy (ICBT) for depressive symptoms. *Internet Interventions* 3 (2016), 1–7.

[37] Michael Hölzer, Erhard Mergenthaler, Dan Pokorny, Horst Kächele, and Lester Luborsky. 1996. Vocabulary measures for the evaluation of therapy outcome: Re-studying transcripts from the Penn Psychotherapy Project. *Psychotherapy Research* 6, 2 (1996), 95–108.

[38] Kurt Hornik, Bettina Grün, and Michael Hahsler. 2005. arules-A computational environment for mining association rules and frequent item sets. *Journal of Statistical Software* 14, 15 (2005), 1–25.

[39] Spencer L James, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi, Hedayat Abbastabar, Foad Abd-Allah, Jemal Abdela, Ahmed Abdelalim, and others. 2018. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 392, 10159 (2018), 1789–1858.

[40] Robert Johansson, Gerhard Andersson, Ebmeier, Smit, Kessler, Cuijpers, Cuijpers, Andersson, Andersson, Andersson, and others. 2012a. Internet-based psychological treatments for depression. *Expert review of neurotherapeutics* 12, 7 (2012), 861–870.

[41] Robert Johansson, Elin Sjöberg, Magnus Sjögren, Erik Johnsson, Per Carlbring, Therese Andersson, Andréas Rousseau, and Gerhard Andersson. 2012b. Tailored vs. standardized internet-based cognitive behavior therapy for depression and comorbid symptoms: a randomized controlled trial. *PloS one* 7, 5 (2012), e36905.

[42] Michael I Jordan and Tom M Mitchell. 2015. Machine learning: Trends, perspectives, and prospects. *Science* 349, 6245 (2015), 255–260.

[43] Evangelos Karapanos. 2015. Sustaining user engagement with behavior-change tools. *Interactions* (2015).

[44] Ramakanth Kavuluru, María Ramos-Morales, Tara Holaday, Amanda G Williams, Laura Haye, and Julie Cerel. 2016. Classification of helpful comments on online suicide watch forums. In *Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*. ACM, 32–40.

[45] Britt Klein, Joanna Mitchell, Jo Abbott, Kerrie Shandley, David Austin, Kathryn Gilson, Litza Kiropoulos, Gwenda Cannard, and Tomi Redman. 2010. A therapist-assisted cognitive behavior therapy internet intervention for posttraumatic stress disorder: pre-, post-and 3-month follow-up results from an open trial. *Journal of anxiety disorders* 24, 6 (2010), 635–644.

[46] Adam DI Kramer, Lui Min Oh, and Susan R Fussell. 2006. Using linguistic features to measure presence in computer-mediated communication. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 913–916.

[47] Kurt Kroenke, Robert L Spitzer, and Janet BW Williams. 2001. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine* 16, 9 (2001), 606–613.

[48] Gionet Kylie. 2018. Meet Tess: The Mental Health Chatbot that Thinks like a Therapists. (2018). https://www.theguardian.com/society/2018/apr/25/meet-tess-the-mental-health-chatbot

[49] Tien-Duy B Le and David Lo. 2015. Beyond support and confidence: Exploring interestingness measures for rule-based specification mining. In *2015 IEEE 22nd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*. IEEE, 331–340.

[50] Reeva Lederman, Greg Wadley, John Gleeson, Sarah Bendall, and Mario Álvarez-Jiménez. 2014. Moderated online social therapy: Designing and evaluating technology for mental health. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 1 (2014), 5.

[51] Philip Lindner, Elinor Linderot Olsson, Amanda Johnsson, Mats Dahlin, Gerhard Andersson, and Per Carlbring. 2014. The impact of telephone versus e-mail therapist guidance on treatment outcomes, therapeutic alliance and treatment engagement in Internet-delivered CBT for depression: A randomised pilot trial. *Internet Interventions* 1, 4 (2014), 182–187.

[52] Melanie Lovatt and John Holmes. 2017. Digital phenotyping and sociological perspectives in a Brave New World. *Addiction (Abingdon, England)* 112, 7 (2017), 1286.

[53] Bernd Löwe, Oliver Decker, Stefanie Müller, Elmar Brähler, Dieter Schellberg, Wolfgang Herzog, and Philipp Yorck Herzberg. 2008. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical care* 46, 3 (2008), 266–274.

[54] Colin Martindale. 1975. English Regressive Imagery Dictionary (RID). (1975). https://rdrr.io/cran/lexicon/man/key_regressive_imagery.html

[55] Mark Matthews, Stephen Voida, Saeed Abdullah, Gavin Doherty, Tanzeem Choudhury, Sangha Im, and Geri Gay. 2015. In Situ Design for Mental Illness: Considering the Pathology of Bipolar Disorder in mHealth Design. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '15)*. ACM, New York, NY, USA, 86–97. DOI:http://dx.doi.org/10.1145/2785830.2785866

[56] Susan Michie, Lucy Yardley, Robert West, Kevin Patrick, and Felix Greaves. 2017. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *Journal of medical Internet research* 19, 6 (2017), e232.

[57] Saif M Mohammad and Peter D Turney. 2013. Nrc emotion lexicon. *National Research Council, Canada* (2013).

[58] David Mohr, Pim Cuijpers, and Kenneth Lehman. 2011. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. *Journal of medical Internet research* 13, 1 (2011), e30.

[59] David C Mohr, Stephen M Schueller, Kathryn Noth Tomasino, Susan M Kaiser, Nameyeh Alam, Chris Karr, Jessica L Vergara, Elizabeth L Gray, Mary J Kwasny, and Emily G Lattie. 2019. Comparison of the Effects of Coaching and Receipt of App Recommendations on Depression, Anxiety, and Engagement in the IntelliCare Platform: Factorial Randomized Controlled Trial. *Journal of medical Internet research* 21, 8 (2019), e13609.

[60] Inbal Nahum-Shani, Shawna N Smith, Bonnie J Spring, Linda M Collins, Katie Witkiewitz, Ambuj Tewari, and Susan A Murphy. 2017. 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 (2017), 446–462.

[61] Michelle G Newman, Lauren E Szkodny, Sandra J Llera, and Amy Przeworski. 2011. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? *Clinical psychology review* 31, 1 (2011), 89–103.

[62] Thin Nguyen, Bridianne O'Dea, Mark Larsen, Dinh Phung, Svetha Venkatesh, and Helen Christensen. 2017. Using linguistic and topic analysis to classify sub-groups of online depression communities. *Multimedia tools and applications* 76, 8 (2017), 10653–10676.

[63] Alicia L Nobles, Jeffrey J Glenn, Kamran Kowsari, Bethany A Teachman, and Laura E Barnes. 2018. Identification of imminent suicide risk among young adults using text messages. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 413.

[64] Matthew K Nock, Irving Hwang, Nancy A Sampson, and Ronald C Kessler. 2010. Mental disorders, comorbidity and suicidal behavior: results from the National Comorbidity Survey Replication. *Molecular psychiatry* 15, 8 (2010), 868.

[65] Emma O'Brien. 2018. *Therapist behaviours, the working alliance and clinician experience in iCBT for depression and anxiety*. Ph.D. Dissertation. Trinity College Dublin.

[66] Theodor Chris Panagiotakopoulos, Dimitrios Panagiotis Lyras, Miltos Livaditis, Kyriakos N Sgarbas, George C Anastassopoulos, and Dimitrios K Lymberopoulos. 2010. A contextual data mining approach toward assisting the treatment of anxiety disorders. *IEEE transactions on information technology in biomedicine* 14, 3 (2010), 567–581.

[67] Pablo Paredes, Ran Gilad-Bachrach, Mary Czerwinski, Asta Roseway, Kael Rowan, and Javier Hernandez. 2014. PopTherapy: Coping with stress through pop-culture. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. ICST (Institute for Computer Sciences, Social-Informatics and Technology, 109–117.

[68] Albert Park, Mike Conway, and Annie T Chen. 2018. Examining thematic similarity, difference, and membership in three online mental health communities from Reddit: a text mining and visualization approach. *Computers in human behavior* 78 (2018), 98–112.

[69] Björn Paxling, Susanne Lundgren, Anita Norman, Jonas Almlöv, Per Carlbring, Pim Cuijpers, and Gerhard Andersson. 2013. Therapist behaviours in internet-delivered cognitive behaviour therapy: analyses of e-mail correspondence in the treatment of generalized anxiety disorder. *Behavioural and cognitive psychotherapy* 41, 3 (2013), 280–289.

[70] Ramesh P Perera-Delcourt and Gemma Sharkey. 2019. Patient experience of supported computerized CBT in an inner-city IAPT service: a qualitative study. *the Cognitive Behaviour Therapist* 12 (2019).

[71] John P Pestian, Pawel Matykiewicz, and Jacqueline Grupp-Phelan. 2008. Using natural language processing to classify suicide notes. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*. Association for Computational Linguistics, 96–97.

[72] Forough Poursabzi-Sangdeh, Daniel G Goldstein, Jake M Hofman, Jennifer Wortman Vaughan, and Hanna Wallach. 2018. Manipulating and measuring model interpretability. *arXiv preprint arXiv:1802.07810* (2018).

[73] Moby Project. 2014. Public-domain lexical resources; word lists, thesaurus, hyphenation, pronunciation. (2014). `https://github.com/Hyneman/moby-project`

[74] Yada Pruksachatkun, Sachin R Pendse, and Amit Sharma. 2019. Moments of Change: Analyzing Peer-Based Cognitive Support in Online Mental Health Forums. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 64.

[75] PWP. 2019. Psychological Wellbeing Practitioner. Overview of Role. (2019). `https://www.instituteforapprenticeships.org/apprenticeship-standards/psychological-wellbeing-practitioner/`

[76] Olga Pykhtina, Madeline Balaam, Gavin Wood, Sue Pattison, Ahmed Kharrufa, and Patrick Olivier. 2012. Magic land: the design and evaluation of an interactive tabletop supporting therapeutic play with children. In *Proceedings of the Designing Interactive Systems Conference*. ACM, 136–145.

[77] Mashfiqui Rabbi, Shahid Ali, Tanzeem Choudhury, and Ethan Berke. 2011. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In *Proceedings of the 13th international conference on Ubiquitous computing*. ACM, 385–394.

[78] Shiquan Ren, Hong Lai, Wenjing Tong, Mostafa Aminzadeh, Xuezhang Hou, and Shenghan Lai. 2010. Nonparametric bootstrapping for hierarchical data. *Journal of Applied Statistics* 37, 9 (2010), 1487–1498.

[79] Stefan Rennick-Egglestone, Sarah Knowles, Gill Toms, Penny Bee, Karina Lovell, and Peter Bower. 2016. Health Technologies' In the Wild': Experiences of Engagement with Computerised CBT. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2124–2135.

[80] Robert Reynes, Colin Martindale, and Hartvig Dahl. 1984. Lexical differences between working and resistance sessions in psychoanalysis. *Journal of Clinical Psychology* 40, 3 (1984), 733–737.

[81] Derek Richards, Angel Enrique, and Jorge E. Palacios. 2019. *Internet-delivered Cognitive Behaviour Therapy*.

[82] Derek Richards and Ladislav Timulak. 2012. Client-identified helpful and hindering events in therapist-delivered vs. self-administered online cognitive-behavioural treatments for depression in college students. *Counselling Psychology Quarterly* 25, 3 (2012), 251–262.

[83] Derek Richards, Ladislav Timulak, Emma O'Brien, Claire Hayes, Noemi Vigano, John Sharry, and G Doherty. 2015. A randomized controlled trial of an internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression. *Behaviour research and therapy* 75 (2015), 20–31.

[84] John Rooksby, Alistair Morrison, and Dave Murray-Rust. 2019. Student Perspectives on Digital Phenotyping: The Acceptability of Using Smartphone Data to Assess Mental Health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 425.

[85] Koustuv Saha and Munmun De Choudhury. 2017. Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 92.

[86] Pedro Sanches, Axel Janson, Pavel Karpashevich, Camille Nadal, Chengcheng Qu, Claudia Daudén Roquet, Muhammad Umair, Charles Windlin, Gavin Doherty, Kristina Höök, and Corina Sas. 2019. 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 (CHI '19)*. ACM, New York, NY, USA, Article 245, 17 pages. DOI: http://dx.doi.org/10.1145/3290605.3300475

[87] VC Sánchez-Ortiz, C Munro, D Stahl, J House, H Startup, J Treasure, C Williams, and U Schmidt. 2011. A randomized controlled trial of internet-based cognitive-behavioural therapy for bulimia nervosa or related disorders in a student population. *Psychological Medicine* 41, 2 (2011), 407–417.

[88] Jessica Schroeder, Chelsey Wilkes, Kael Rowan, Arturo Toledo, Ann Paradiso, Mary Czerwinski, Gloria Mark, and Marsha M Linehan. 2018. 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*. ACM, 398.

[89] Stephen M Schueller, Kathryn Noth Tomasino, and David C Mohr. 2017. Integrating human support into behavioral intervention technologies: the efficiency model of support. *Clinical Psychology: Science and Practice* 24, 1 (2017), 27–45.

[90] Eva Sharma and Munmun De Choudhury. 2018. Mental Health Support and its Relationship to Linguistic Accommodation in Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 641.

[91] Adrian BR Shatte, Delyse M Hutchinson, and Samantha J Teague. 2019. Machine learning in mental health: a scoping review of methods and applications. *Psychological medicine* 49, 9 (2019), 1426–1448.

[92] Amit Singhal, Chris Buckley, and Manclar Mitra. 2017. Pivoted document length normalization. In *ACM SIGIR Forum*, Vol. 51. ACM, 176–184.

[93] Michael Stigler and Dan Pokorny. 2001. Emotions and primary process in guided imagery psychotherapy: Computerized text-analytic measures. *Psychotherapy Research* 11, 4 (2001), 415–431.

[94] William B Stiles, Lara Honos-Webb, and Michael Surko. 1998. Responsiveness in psychotherapy. *Clinical psychology: Science and practice* 5, 4 (1998), 439–458.

[95] Anja Thieme, John McCarthy, Paula Johnson, Stephanie Phillips, Jayne Wallace, Siân Lindley, Karim Ladha, Daniel Jackson, Diana Nowacka, Ashur Rafiev, and others. 2016. Challenges for designing new technology for health and wellbeing in a complex mental healthcare context. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2136–2149.

[96] Nickolai Titov. 2011. Internet-delivered psychotherapy for depression in adults. *Current opinion in psychiatry* 24, 1 (2011), 18–23.

[97] Nickolai Titov, Blake F Dear, Genevieve Schwencke, Gavin Andrews, Luke Johnston, Michelle G Craske, and Peter McEvoy. 2011. Transdiagnostic internet treatment for anxiety and depression: a randomised controlled trial. *Behaviour research and therapy* 49, 8 (2011), 441–452.

[98] Truyen Tran, Dinh Phung, Wei Luo, Richard Harvey, Michael Berk, and Svetha Venkatesh. 2013. An integrated framework for suicide risk prediction. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1410–1418.

[99] Daniel Vigo, Graham Thornicroft, and Rifat Atun. 2016. Estimating the true global burden of mental illness. *The Lancet Psychiatry* 3, 2 (2016), 171–178.

[100] Birgit Wagner, Andrea B Horn, and Andreas Maercker. 2014. Internet-based versus face-to-face cognitive-behavioral intervention for depression: a randomized controlled non-inferiority trial. *Journal of affective disorders* 152 (2014), 113–121.

[101] Rui Wang, Fanglin Chen, Zhenyu Chen, Tianxing Li, Gabriella Harari, Stefanie Tignor, Xia Zhou, Dror Ben-Zeev, and Andrew T Campbell. 2014. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*. ACM, 3–14.

[102] Harvey A Whiteford, Louisa Degenhardt, Jürgen Rehm, Amanda J Baxter, Alize J Ferrari, Holly E Erskine, Fiona J Charlson, Rosana E Norman, Abraham D Flaxman, Nicole Johns, and others. 2013. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *The Lancet* 382, 9904 (2013), 1575–1586.

[103] Jesse H Wright, Jesse J Owen, Derek Richards, Tracy D Eells, Thomas Richardson, Gregory K Brown, Marna Barrett, Mary Ann Rasku, Geneva Polser, and Michael E Thase. 2019. Computer-Assisted Cognitive-Behavior Therapy for Depression: A Systematic Review and Meta-Analysis. *The Journal of clinical psychiatry* 80, 2 (2019).

[104] Xuhai Xu, Prerna Chikersal, Afsaneh Doryab, Daniella K Villalba, Janine M Dutcher, Michael J Tumminia, Tim Althoff, Sheldon Cohen, Kasey G Creswell, J David Creswell, and others. 2019. Leveraging Routine Behavior and Contextually-Filtered Features for Depression Detection among College Students. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 116.

[105] Amir Hossein Yazdavar, Hussein S Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-supervised approach to monitoring clinical depressive symptoms in social media. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*. ACM, 1191–1198.

[106] Farzaneh Zahedi and Mohammad-Reza Zare-Mirakabad. 2014. Employing data mining to explore association rules in drug addicts. *Journal of AI and Data Mining* 2, 2 (2014), 135–139.