Exploring the Effects of Technological Writing Assistance for Support Providers in Online Mental Health Community

Zhenhui Peng, Qingyu Guo, Ka Wing Tsang, Xiaojuan Ma
The Hong Kong University of Science and Technology, Hong Kong
{zpengab, qguoag, kwtsangae}@connect.ust.hk, mxj@cse.ust.hk

ABSTRACT
Textual comments from peers with informational and emotional support are beneficial to members of online mental health communities (OMHCs). However, many comments are not of high quality in reality. Writing support technologies that assess (AS) the text or recommend (RE) writing examples on the fly could potentially help support providers to improve the quality of their comments. However, how providers perceive and work with such technologies are under-investigated. In this paper, we present a technological prototype MepsBot which offers providers in-situ writing assistance in either AS or RE mode. Results of a mixed-design study with 30 participants show that both types of MepsBots improve users' confidence in and satisfaction with their comments. The AS-mode MepsBot encourages users to refine expressions and is deemed easier to use, while the RE-mode one stimulates more support-related content re-editions. We report concerns on MepsBot and propose design considerations for writing support technologies in OMHCs.

Author Keywords
Mental health; online community; writing support tools; informational support; emotional support

CCS Concepts
•**Human-centered computing** → **Empirical studies in HCI;** *User interface toolkits;*

INTRODUCTION
Online mental health communities (OMHCs) offer a promising way for people with mental issues to access peer support – "providing knowledge, experience, emotional, social, or practical help to each other" [47, 51]. For example, in Reddit *r/depression* – an anonymous OMHC with over 493k members up to March 2019 [62], one can create a support-seeking post when suffering from depression, and others can respond by text under the post. The success of OMHC relies largely on the exchange of informational (e.g., advice) and emotional (e.g., empathy) support among community members [37, 44, 66]. However, in reality, many peer comments are not of high quality, lacking practical information or being disrespectful [44, 51, 52]. Such comments could distress the support seekers, resulting in their withdrawal from OMHCs and further isolation [44, 52]. Meanwhile, peers (especially newcomers [6, 75]) who fail to provide meaningful support may experience a reduced sense of worthiness and competence with a likely decrease of commitment to the community [9, 51, 75].

Existing OMHCs have experimented with various methods to improve the quality of informational and emotional support offered by peers. One common approach is to flag bad comments either by community members (e.g., downvote option on Reddit), moderators [26], or algorithms [14]. This may filter out deliberately hurtful messages; however, for those who mean to help, it is a kind of delayed feedback that may impair their confidence, interest, and reputation [41, 48, 76]. Another means to safeguard peer support quality is to train providers beforehand, such as teaching them psychotherapeutic techniques (e.g., cognitive reappraisal) online [31, 42, 49]. However, this method usually requires weeks of effort from support providers and suffers from a high attrition rate [4, 31]. In other words, the two approaches above do not provide timely assistance for peers to manifest informative and emotionally caring responses to those in need.

A more plausible alternative is to offer writing feedback and guidance on the fly to the support providers with technological aids, which can take various forms and has been investigated in other domains [32, 70, 74]. For example, Wu et al. designed the *Additional Writing Help* tool to assess Facebook posts drafted by users with dyslexia before publishing and suggest words for refinement [74]. Their tool is shown to help dyslexic users express themselves more confidently [74]. However, assessing writing performance might discourage support providers who are also OMHC members with mental health issues, as they are sensitive to being judged, either by audience members or algorithms [24, 30, 52]. Instead of directly judging the content, there exists another type of writing support tool that recommends examples related to the input text as references for improvement [32, 40]. For instance, Gero et al. proposed *Metaphoria* that generates coherent metaphorical examples based on an input word to promote creative writing like poetry [32]. Such recommendation tools could generally guide users to write more expressively but are known to raise sincerity and authorship concerns [32, 40]. This may

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

ACM. ISBN 978-1-4503-6708-0/20/04. . . $15.00
DOI: https://doi.org/10.1145/3313831.3376695

become an issue in OMHCs, where authentic self-presentation is highly appreciated [42]. Hence, despite the success of writing support tools in other domains, there is the need to explore how the support providers' commenting process and outcome are affected by such tools, how they perceive and work with different types of assisting mechanisms, and what the design considerations are for such technologies in OMHCs.

To this end, we design a technological prototype called MepsBot (**Me**ntal health **p**eer **s**upport **Bot**) to explore user perception and behavior given various types of in-situ assistance. Specifically, we first use well-developed techniques to model the comment's levels of informational support (IS) and emotional support (ES), two types of social support that have been found theoretically and empirically critical in mental health support [34]. Then we design two modes of MepsBot to offer assistance, i.e., 1) assessment (AS): scoring IS and ES of the input text and eliciting feedback on how to improve; and 2) recommendation (RE): suggesting high-IS and high-ES example comments that are semantically similar to the current draft with key features highlighted as a reference for revision. To evaluate the effects of MepsBot on support providers' perception and behavior as well as on the outcome of the comments in OMHCs, we conduct a mixed-method, 2 (within-subjects: time pre vs. post) by 2 (between-subjects: mode AS vs. RE) mixed-design study. We invite 30 participants who have experienced or are combating depression and are willing to help others in similar situations online. The results show that both types of MepsBot improve participants' confidence in the IS and ES manifested in their comments. The AS-mode MepsBot motivates participants to refine their expressions and is perceived as easier to use. The RE-mode MepsBot stimulates users to re-edit the IS- and ES-related contents in their comments. Our work adds to the understanding on how support providers perceive and work with the on-the-fly technological writing assistants in OMHCs. We also propose the design considerations for such technologies about how to improve user experience and address the concerns.

RELATED WORK

Peer Support in Online Mental Health Communities

With the benefits of anonymity, empowerment, and easy access, online mental health communities (OMHCs) have emerged as valuable resources for peer support [35, 51]. OMHC members can exchange informational support (IS) and emotional support (ES) via posting threads [19, 37, 44]. The former takes the form of "advice, knowledge, and information", while the latter includes "understanding, encouragement, affirmation, sympathy, or caring" [67, 72]. High-quality peer support circulated inside OMHCs not only benefits support seekers by enhancing social connectedness [51] and evoking positive cognitive change [58], but also does good to support providers by enhancing self-esteem (e.g., feeling worthwhile) [9] and acquiring social skills [65]. Unfortunately, many comments in OMHCs are of relatively low IS and ES [68]. Based on the results of expert annotation and model classification, Sharma et al. found that over 67% of comments across assorted Reddit OMHCs failed to convey high IS or ES in general [68]. The low-quality comments could be ascribed to deliberately unsupportive members [44, 52]. They may also be made by peers who intend to be supportive but have difficulties (e.g., being uncertain of the social norms or their own expertise [3]) expressing their support [18]. Our research is motivated by the benefits of high-quality peer support, and we try to facilitate providers to manifest it in OMHCs.

Quality Assurance of Peer Support Interactions in OMHCs

Existing OMHCs have explored different ways to encourage high-quality support exchanged among peers. The first category of means is content moderation, with the social voting system being one of the most common methods. For example, Reddit allows users to "downvote" those comments that contribute little to the conversation, which directly affects the posters' "karma" points – an indication of reputation. In this way, members are likely to be more careful about what they write [39, 77]. Apart from peer voting, many OMHCs involve human moderators [26] or automatic algorithms to scrutinize content. The latter has gained increasing popularity owing to the recent advances in artificial intelligence, such as detection of negative content that violates community rules [13, 17] and prediction of seekers' needs and comments' effectiveness [15, 16, 21, 58]. While content moderation can detect less qualified support, there is usually a lack of reciprocal feedback to the providers regarding what they have done wrong. Even though Reddit encourages viewers to attach an explanation for downvotes [61], it is not guaranteed that such constructive criticism is always available. As a result, these content moderation methods can silence those who may want to express themselves without fear of judgment [30], and may reduce people's interest and confidence in providing support [41, 48].

The second category of approaches to boost support quality in OMHCs is teaching members how to exercise psychotherapeutic techniques when offering help. For instance, Lederma et al. built a clinician-moderated site for a group of young people suffering First-Episode Psychosis to learn positive psychology [42]. They showed that it was an effective and engaging way to build supportive accountability towards other members on the site through a field study [42]. Alvarez et al. developed a similar online instructional tool that involves professionals in the education of peer supporters with schizophrenia [28]. They found the tool beneficial, as 90% of the users reported wanting to become online peer moderators after the training. Morris et al. designed a peer-to-peer cognitive reappraisal platform to promote evidence-based techniques, and found that peers inside significantly improve the abilities of reappraisal after a three-weeks trial [49]. Despite the effectiveness of these platforms, users usually need weeks of effort to learn, and they could drop out of the courses as the treatment may not be tailored to their needs and interests [4, 25]. Moreover, they do not offer in-situ assistance to peers in providing support.

As an alternative, Leary et al. proposed to guide peers by fixed prompts on how to express and reflect thoughts using a problem-solving framework, which leads to "deep" chats with solutions to problems and new perspectives [53]. In the discussion, they suggested that the guidance should be adaptive in the context of the ongoing conversation. Building upon this finding, our work explores the efficacy of on-the-fly, adaptive writing assistance to peers when they are composing comments to provide support in OMHCs.

Writing with Technological Assistance

Researchers in Human-Computer Interaction (HCI) have investigated different ways to facilitate people's writing experiences in various domains. One popular type of mechanisms, noted as assessment-based methods, usually presents authors with a direct assessment of input text regarding certain characteristics or expected effects of interests [29, 38, 69, 70, 74]. For example, to promote positive reflection, Wang et al. designed MirrorU that captures emotional components in users' daily reflective writings in real time on mobile devices [70]. It then displays timely visual and textual feedback next to the writings to nudge users to compose longer reflections with more positive emotional words. To encourage effort in close interpersonal communication, Kelly et al. designed Message Builder that keeps track of the word count to induce longer social messages than users' previous written ones [38]. Also, to help users be more cautious about personal information before posting on Facebook, Wang et al. developed an interface that sets a timer and presents possible audiences of the post [71].

Another kind of mechanisms, noted as recommendation-based methods, advocates learning from examples [32, 36, 40]. For instance, Kim et al. proposed Lily, which recommends romantic lyrics similar to the current message as people type, and showed that it could facilitate affectionate communication between couples [40]. To support writing introductory help requests in emails, Hui et al. presented IntroAssist, which provides tagged peer examples for the novice entrepreneurs, and showed that the tool could help the novices write requests more effectively [36]. To facilitate short story writing, Clark et al. proposed a prototype that suggests the next sentence based on the context provided by users, and the writing process with their system is perceived to be fun and helpful [20].

Despite their benefits, assessment-based writing tools could frustrate users [38, 71], while the recommendation-based systems could raise concerns about authorship and sincerity [32, 40]. These issues may be even more severe in an OMHC setting given that members experiencing mental challenges are sensitive to judgement and insincerity [30, 42]. It is thus worthwhile to examine user behavior and perception towards these two types of assisting methods in the context of generating support in OMHCs. We are particularly interested in gaining insights into the design of proper technology for promoting high-quality IS and ES among peers.

MEPSBOT SYSTEM

In this section, we present MepsBot, a prototype system for assisting provider's manifestation of peer support in OMHCs, to explore the following research questions:

RQ1: How would the writing support system and its assisting mechanisms affect a) providers' perception and behavior in the process of writing comments in response to support-seeking posts in OMHCs and b) the outcome of comments?

RQ2: How would providers perceive writing support tools for peer support in OMHCs, and what are their concerns?

Note that we do not intend MepsBot to be a canonical technology for peer support assistance, but rather as a tool to probe user experience under different writing support mechanisms to gain plausible design insights. To make the feedback and guidance potentially more engaging [64], we frame the system as a bot to present its messages. In the rest of this section, we first present a data-driven method to build up a support evaluation model, and then detail the design of MepsBot's assessment (**AS**) mode and recommendation (**RE**) mode.

Support Evaluation

To evaluate the levels of IS and ES in textual comments, we apply a well-developed, data-driven approach to build up support classification models [58, 68, 76]. As the content of comments could vary across different OMHCs [68], we choose the Reddit *r/depression* community – a supportive space for anyone struggling with depression (a common mental disorder worldwide [54]) – as a case to demonstrate the usage of MepsBot. We use Pushshift API [59] to retrieve publicly available posts and their comments in *r/depression*. Given a large amount of data available (*r/depression* has over 493k members up to March 2019), we randomly sample 5% posts during 2017.03.01 - 2019.03.01 and collect their comments to ensure the computational efficiency of MepsBot. We exclude the comments from the original authors of the posts and those that are less than three words. This results in 48,148 comments in our dataset. Following the method in [68], we randomly select 450 comments from the dataset and label the level (1 - low, 2 - medium, 3 - high) of IS and ES in each entry. For IS we refer to the occurrence of "advice, referrals or knowledge" [72], while for ES we look for evidence of "understanding, encouragement, affirmation, sympathy, or caring" [72]. Two researchers familiar with the *r/depression* community first rate 25 random samples separately. Then they meet with a third researcher who holds a mental health first aid certificate issued by MHFA [7] to discuss and come up with a consistent rating scheme. After that, the two researchers proceed to rate the remaining comments independently. The inter-rater metric Cohen's κ of initial 450 ratings are 0.860 for IS and 0.892 for ES. The disagreement is resolved by the third researcher. Finally, we have 199 (127), 149 (197), 102 (126) comments labeled as low, medium, and high in IS (ES), respectively.

To train the support classification models, we compile a set of features commonly employed by works on health community content analysis [21, 58, 72, 76] to represent each comment, including 64 features from the LIWC 2015 dictionary [60], a binary feature corresponding to the presence of URL(s), word count (weighted by 60), and sentence count (weighted by 10) [1]. We experiment with various machine learning models – SVM, Multinomial Logistic Regression, Random Forest (RF) and XGBoost. Among them, RF achieves the best result in predicting IS, while XGBoost performs the best in predicting ES, both in terms of mean accuracy and mean F1 score in 10-fold cross-validation (Table 1). Specifically, IS classifier (RF)

[1] The LIWC features and weights are chosen based on trial and error. If word count ≤ 60, then this feature is 1; same as sentence count.

	A	P	R	F1	Top-6 important features
IS	.64 /.07	.64 /.11	.62 /.07	.62 /.07	Word count(.066), I(.048), Sentence count(.033) Social(.031), Pronouns(.027), Personal Pronouns(.024)
ES	.68 /.11	.70 /.14	.68 /.12	.68 /.11	Word count(.090), Sentence count(.045), Personal Pronouns(.028), I(.027), Positive emotions(.026), You(.022)

Table 1. Performance metrics (Accuracy, Precision, Recall, F1-score) and the top-6 important features for IS and ES classifiers. The metrics (μ/σ) are reported based on 10-fold cross-validation. The importance is measured by the proportion of times each feature is used to split the data across all tress in the Random Forest or XGBoost.

Features	IS (1/2/3)	ES (1/2/3)	Scripts sample in AS mode
1) Length-based (Sentence/ Word count)	.45 (.30)/ .68 (.30)/ .90 (.21)	.41 (.27)/ .59 (.32)/ .90 (.20)	"Try write down more stuffs."
2) Personal pronouns (e.g., I, you)	.15 (.08)/ .12 (.06)/ .11 (.04)	.10 (.07)/ .15 (.08)/ .14 (.05)	(IS) "Maybe you can share some knowledge." / (ES) " Show connections to the help seeker, using words like:"
3) Social (e.g., friend, family)	.10 (.10)/ .11 (.08)/ .13 (.06)	-	"Share experience about yourself or others, like:"
4) Positive emotion (e.g., exciting, brave)	-	.03 (.04)/ .05 (.07)/ .05 (.03)	"More positive words can be used in the comment, like:"

Table 2. Four types of important features in predicting IS or ES, and their distributions (mean (SD)) in each level (1-low/2-medium/3-high) of IS and ES predicted by the classifiers on the labeled 450 comments. Feature 1 measures word count (weighted by 60). Features 2, 3, and 4 measure the percentage of corresponding feature words in the comment. Scripts sample is AS-mode MepsBot's utterance to give suggestions.

achieves a mean accuracy of 64% (vs. the best baseline 62%), and the ES classifier (XGBoost) gets a mean accuracy of 68% (vs. the best baseline 67%). We use these two classifiers to label our entire dataset of 48,148 comments, and randomly sample 100 comments to conduct a cross-verification with ground truth labels provided by one of the two annotators. Our models achieve a 65% accuracy for IS and 75% for ES in this exercise, indicating their robustness in performance.

To understand which features could be essential to high IS and ES, we first examine the top-6 most important features of the trained models. Similar to [58], we measure the importance of features as the normalized number of times (note: the larger, the more predictive) a feature is used to split data across all trees in each classifier (Table 1). We then summarize these features into four types, and we calculate their distributions in each level of IS and ES predicted by the classifiers on the labeled 450 comments (Table 2): **1)** Length-based: includes the word count and sentence count; important for both IS and ES predictions; more words could indicate higher IS and ES. **2)** "Personal pronouns": under category "pronouns"; includes categories "I" (e.g., I'll, I mean), "we" (e.g., let's, us), "you" (e.g., your, you've), "shehe" (e.g., herself, he'd), "they" (e.g., their, they are); important for both IS and ES predictions; a higher percentage of "personal pronouns" words could indicate lower IS but higher ES (medium or high). **3)** "Social": includes categories "friend" (e.g., best friend, classmate) and "family" (e.g., brother, father); important for IS prediction; a higher percentage of "social" words could indicate higher IS. **4)** "Positive emotion" (e.g., exciting, brave): important for ES prediction; a higher percentage of "positive emotion" words could indicate higher ES (medium or high). We embed the summarized features (Table 2) into the MepsBot's design.

Figure 1. Design of MepsBot. (a) Assessment mode: (i) report IS and ES scores and (ii) give feedback for improvement. (b) Recommendation mode: (iii)&(iv) suggest example comments that could be semantically similar to user's writing, with highlighted feature words.

Assessment (AS) Mode

MepsBot in its assessment (AS) mode detects and displays the levels of IS and ES of an input draft, and it suggests ways to improve (following the practice of AS-based systems [38, 70]). To be more specific, after a brief self-introduction to inform its ability, MepsBot reports the predicted IS and ES levels of the drafted comment and suggests a point in the current writing (listed in Table 2) that the user could further improve (Figure 1a). We only suggest one point at a time so as not to overwhelm the user, and we choose the point based on the logic flow depicted in Figure 2. We have several considerations in deriving this flow: **1)** To reflect current writing performance, MepsBot has different utterances to suggest possible improvement based on the IS and ES scores [70]. **2)** As the repetitive emphasis on one text feature (especially word count) could frustrate users [38], MepsBot only brings up the length-based issue at the first occurrence of such issue. Also, it has a random factor when choosing among features 2, 3, and 4 (Table 2) to cover all potential issues rather than always promotes the most prominent one. **3)** In promoting the feature 2, 3, or 4, MepsBot randomly suggests twelve related words from the LIWC 2015 dictionary to induce serendipity for writing inspiration [5] [2]. We prepare multiple scripts for feedback on each feature (script samples in Table 2).

Recommendation (RE) Mode

MepsBot in its recommendation (RE) mode suggests high-quality example comments that are semantically similar to the user's current draft from a recommendation pool (following the practice of existing RE-based systems [32, 40]). The recommendation pool consists of 9,080 comments with both high IS and high ES scores labeled by our support classifiers

[2]Open-source in https://github.com/PenguinZhou/MepsBot

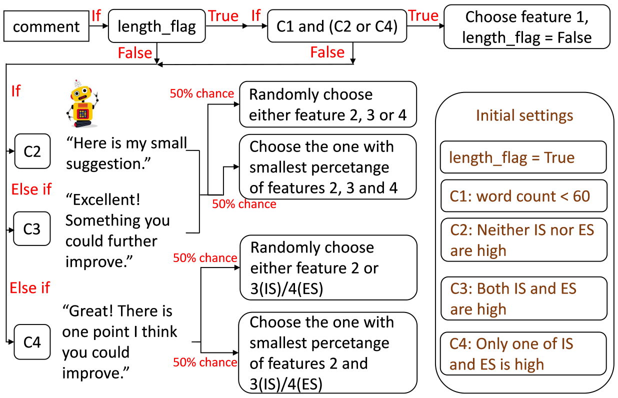

Figure 2. Logic flow for AS-mode MepsBot to suggest improvement on which feature (as listed in Table 2). MepsBot has different scripts in conditions 2,3 and 4. To mitigate potential frustration of continuous emphasis on one feature, it promotes feature 1 only once in a commenting procedure and has a random factor (50% chance) in C2, C3 and C4.

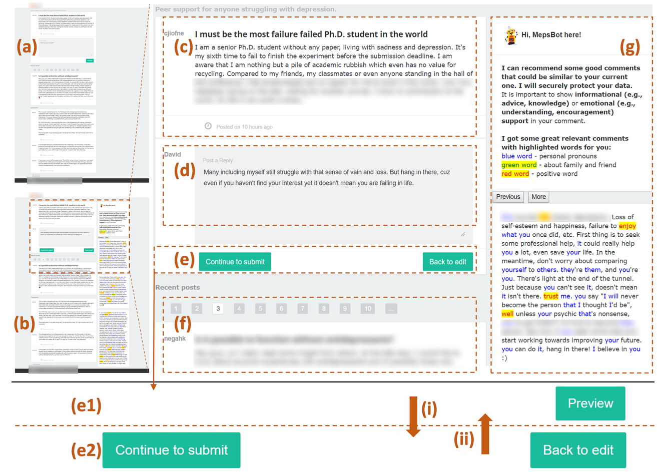

Figure 3. Screenshots of the simulated community: (a) original full page; (b) full page with the RE-mode MepsBot; (c) the target support-seeking post; (d) the text box for writing a comment; (e) buttons of the comment flow, e1 $\leftrightarrow$ e2 if users click (i) "Preview" or (ii) "Back to edit"; (f) recent posts and their comments; (g) MepsBot shows up with the first click on "Preview" and keeps visible until the click on "Continue to submit".

from the original 48,148 comments. We do not suggest comments based on the original support-seeking post to ensure that the suggested examples are more coherent to the provider's thought and are diverse in expressions [1, 32]. We follow the two-step method employed in [50] to retrieve candidate comments, which could enable MepsBot to respond in near real time ($\approx$ 0.9s in our case) when finding semantically similar comments. Specifically, with the input comment, MepsBot first applies Elasticsearch python API [27] to conduct a *More Like This Query* in the pool. This query identifies the most representative terms of the input comment based on tf-idf (term frequency-inverse document frequency), and then returns 50 potentially relevant comments containing such terms as quickly as possible. MepsBot further measures the similarity between the user comment and each of the 50 candidate comments using cosine similarity, which can find semantically similar comments even they only have few words in common [40]. To do so, it encodes every comment into a 768-dimension vector using the BERT-Based, uncased language model, which has state-of-the-art performances in semantic understanding tasks [22] at the time of MepsBot's development. This step returns the top-18 candidates based on the ranking on cosine distance in the vector space. MepsBot's messages and the 18 suggested high-quality example comments are displayed as seen in Figure 1b, for the following considerations: **1)** MepsBot starts with an introduction to show its ability. **2)** To facilitate the visual search for linguistic features, the words that belong to "personal pronouns", "social" or "positive emotion" are marked in different colors [36]. **3)** To not overwhelm the user, each page only shows three examples (note: user can scroll up/down the page), and there are "Previous" and "More" buttons for users to navigate the list of examples.

EXPERIMENT

To investigate MepsBot's impact on support providers' experience in the commenting process and the outcome of that process, we conduct a mixed-method, 2 (within-subjects factor, time: pre vs. post) by 2 (between-subjects factor, mode: AS vs. RE) mixed-design lab study with 30 participants.

Simulated OMHC Environment and MepsBot Setup

We simulate an anonymous online peer support community (Figure 3a) for people who are depressed or struggling with depression. As shown in Figure 3(c)(d)(f), the community website has fundamental elements similar to those in *r/depression*, such as a target help-seeking post, a text box below it for entering a comment, as well as other recent posts and their comments (randomly sampled from *r/depression*). To embed MepsBot into the community, we adopt the setup used in [71], and modify the interaction flow as: preview $\leftrightarrow$ (back to edit) $\rightarrow$ continue to submit (Figure 3e). Each time providers write the very first draft of their comments, the MepsBot is hidden with only a "Preview" button indicating the pathway to the next step. To avoid interrupting users' composing process [74], MepsBot only shows up after their first click on "Preview" (Figure 3g), and then the buttons in part (e) change from (e1) to (e2). Users can choose "Back to edit" (buttons changed from e2 back to e1) to adjust their comments or "Continue to submit" to formally publish the content whenever they want. MepsBot updates its assessment or recommendation on the current draft every time users click "Preview" and hides again upon comment submission. The simulated community is hosted on a server deployed on a 15.6-inch laptop with an Intel i7-7700HQ CPU. It takes around 0.1s/0.9s for AS-/RE-mode MepsBot to respond to a user-triggered preview command.

Participants

We recruit thirty students (13 **F**emales, 15 **M**ales, 2 **N**ot **A**vailable; age range 20-30, $M = 24.37, SD = 2.72$) from a local university via word of mouth. The inclusion criteria are that participants have prior experiences of receiving/providing mental support from/to others and that they are willing to offer support to peers in online communities. All participants are fluent in reading and writing in English (TOEFL read/write $\geq$ 22 or IELTS read/write $\geq$ 6.5). Six of them are undergraduates, and the rest are postgraduates majoring in a variety of fields (e.g., computer science, economics). Previous research suggests that the majority of college students experience some forms of mental distress [8, 46, 63]. The self-assessments on the severity of depression (measure: PHQ-9 [45]) in pre-screening show that our participants could have at least moderate depression: seven moderate (score 10-14), 18 moderately severe (15-19), and five severe (20-27) [53]. The depressing issues they encountered are mostly about Ph.D. research (count: 15), relationships with boy/girlfriend or families (6), and future direction (4). Six of them report having experience in OMHCs, and all participants are very familiar with online communities. We randomly assign the participants to the AS group (AP1-15) or the RE group (RP1-15).

Tasks

In the experiment, each participant is asked to complete three tasks. Each task involves writing a comment for one of three given posts from peers to show their support. As the same ideologies among members are important for engagement in OMHCs [3], we invite another three postgraduates (Males, age: 24, 24, 22) as the support seekers to each create an anonymous post about their recent depressing issue. They are all suffering from moderately severe depression and have the desire to communicate their issues in OMHCs. The topics of the three posts are Ph.D. research failure, restless working under financial pressure from girlfriend and family, and the loss of future goals. We counterbalance the order of presenting these posts to each participant to mitigate potential order effects.

Hypotheses and Mixed-Method Measurements

Providers' perception of and behavior in the commenting process, and the outcome of comments (RQ1)

To evaluate the effects of MepsBot, we define two timestamps in a commenting procedure of each task in our study:

Pre-: the moment right before a user's very first click on the "Preview" button in a task – the user writes a comment (denoted as pre-comment) without MepsBot's assistance.

Post-: the moment when a user clicks the "Continue to submit" button – the user finalizes a comment (post-comment) after getting assessments or recommendations from MepsBot.

Perception. Previous works suggest that writing support technologies could improve users' confidence and satisfaction when posting on social media [74]. In particular, it is suggested that providing assessment offers more direct feedback on writing performance [70] than showing examples. Therefore, we hypothesize that: ***H1.*** Overall, compared to the pre-comments, participants are significantly more confident that their post-comments can provide *(H1a)* IS and *(H1b)* ES to the support seekers, and *(H1c)* they are significantly more satisfied with their post-comments. *(H1d)* The ratings on the three measures of post-comments are significantly higher in the AS group than those in the RE group. We adapt the standard 7-point Likert scale metrics from [74] (e.g., 1 not at all confident/satisfied - 7 extremely confident/satisfied).

Behavior. To inspect participants' behavior during the commenting process with MepsBot, we log the following data in each task session: 1) task completion time – from opening the target post to clicking on "Continue to submit"; 2) the number of times that a user triggers "Back to edit"; 3) the draft of the comment each time a user presses "Preview" so as to check whether the user actually changes the content or not.

Outcome of comments. To understand whether and how MepsBot affects the outcome of a commenting process, we track how the content of the post-comments is changed (if any) from pre-comments in the AS and RE groups: 1) the word counts of pre- and post-comments; 2) the levels of IS and ES in pre- and post-comments predicted by MepsBot's support classifiers; 3) if any, the types of change between every two consecutive versions of previewed comments in each task session.

Providers' perception of and concerns about MepsBot (RQ2)

Acceptance and user experience. Considering that AS-mode MepsBot provides direct feedback while the RE-mode MepsBot has concrete examples, we hypothesize that: ***H2.*** Compared to RE mode, MepsBot in AS mode is perceived significantly *(H2a)* easier to use, *(H2b)* more likeable, but *(H2c)* less useful in helping users to write comments to support others online. We adapt metrics from [10, 43, 56] to measure each aspect with one item on a standard 7-point Likert scale.

Concerns. To elicit users' potential concerns with the accuracy and the use of MepsBot, we ask participants to rate: 1) perceived accuracy of its assessment or recommendation [74] (indicating a number in the range of 0 - 100% using a slide bar); 2) one item for overall perceived risks, and one item for privacy concern [73] (1 strongly disagree - 7 strongly agree). To better understand the reasons behind the aforementioned data, we conduct a semi-structured interview with each participant at the end of the experiment. The interview starts with participants reporting their overall impressions of MepsBot, including its pros and cons. We then ask about their thoughts for each rating. Regarding *concerns*, we specifically ask if they are frustrated by AS-mode MepsBot's judgment [30] or if they worry about the sincerity of comments with RE-mode MepsBot [40] to verify consistency with previous literature. Finally, we collect their suggestions for MepsBot.

Procedure

After obtaining participants' consent, we first showcase the simulated community and tell them it is a mapping from a real one. We then let them freely explore it for three minutes. Next, we introduce the task and walk the participants through the MepsBot system. In each of the three task sessions, the participants are asked to write a comment in response to a given support-seeking post. They can "Preview" or go "Back to edit" the comment whenever and as many times as they want. After the participants hit the "Continue to submit" button, we ask them to fill out a survey to rate perceptions of the commenting process. Upon completion of all the three tasks, we instruct the participants to finish one final questionnaire about the perception of MepsBot and potential concerns. Then we conduct a semi-structured interview with them to make sense of the ratings. After debriefing, we give the participants some compensations and chat with them to ensure that they leave in a stable emotional status. The whole study lasts for 40-60 minutes for each participant.

Ethical Issues and Safety Protocol

Prior to this work, we obtained IRB approval for broader research project on patients' and caregivers' practices of healthcare service systems and online communities, covering both our data collection and analysis as well as interviews with participants. When citing the anonymous comments from *r/depression* and from our participants, we remove all identifiable information in the text to protect their privacy. In addition, we set up a protocol to ensure the safety of our support seekers and providers during the study. For the invited three seekers, we contact them every day during the course of the study to check if they are mentally stable. For the 30 providers, we allow them to exit the study with full pay at any point of time if they feel uncomfortable and suggest them consult the Counseling and Wellness Center (contact provided) in this condition. All of them finish the study successfully.

DATA ANALYSIS

To inspect the possible effects of post order on our statistical measurements, we first run a set of mixed ANOVA (mode and post order as between-subjects factors) tests on each of the quantitative measures (as within-subjects) of participants' perceptions and behaviors. Neither the main effect of post order nor its interaction effects are significant. Therefore, in the following statistical analysis, we treat MepsBot's mode as the only between-subjects variable. To test ***H1***, we conduct a two-way mixed ANOVA on the repeated measures (as within-subjects) about pre-comment and post-comment in each session. For the measures in ***H2*** between two modes, we run the independent-samples t-test.

As for the changes between two continuously previewed comments in a task session, we conduct a thematic analysis [11] to check what types of content are changed. We collect the sentences that are added, removed, or modified in the re-edition. Two authors first code all the changed parts in these sentences independently; after rounds of comparison and discussions, we agree on a starting list of codes. After iterative coding, we agree that the changes on IS-, ES-related content, and expression are overarching themes. We name the themes and categories of the codes, and count their occurrences in AS and RE groups. We apply the same thematic analysis to the transcribed interview data to understand the users' reasons for their perception and behavior.

RESULT

In this section, we present our findings of providers' perceptions and behaviors in the commenting process, the outcome of comments, and feelings towards MepsBot.

Perception and Behaviors in the Process (RQ1a)

Perception. We compare participants' overall perceptions on pre- and post-comments (Figure 4), and find them significantly more confident of their ability to provide IS to the support seekers after receiving help from MepsBot ($M = 5.11, SD = 1.28$) than before ($M = 4.42, SD = 1.16$); mixed ANOVA, $F(1,88) = 35.4, p < .01, \eta^2 = .29$; ***H1a*** accepted. Their confidence in providing ES also improves significantly from pre- ($M = 4.33, SD = 1.40$) to post-comments ($M = 5.11, SD = 1.50$); $F(1,88) = 42.2, p < .01, \eta^2 = .32$; ***H1b*** accepted. In addition, they are significantly more satisfied with their post- ($M = 5.63, SD = 0.91$) than with pre-comments ($M = 4.83, SD = 1.14$); $F(1,88) = 63.9, p < .01, \eta^2 = .42$; ***H1c*** accepted. The average confidence (IS and ES) and satisfaction ratings under each mode are higher in post-comments than those in pre-comments as well (as depicted in Figure 4). There is no significant interaction effect between such perception changes and the mode of MepsBot on these three measures, but the between-subjects effects on them are all significant ($p < .05$). The follow-up independent-samples t-tests (between-subjects factor: AS vs. RE) on the three measures of post-comments show that participants in the AS group generally have significantly higher confidence in the IS ($t_{88} = 2.93, p < .01$) and ES ($t_{88} = 2.00, p < .05$) provided in their post-comments and are significantly more satisfied with their output ($t_{88} = 3.32, p < .01$) than those in the RE group; ***H1d*** accepted. Eight participants (6 in AS, 2 in RE) explicitly mention how MepsBot boosts their confidence and satisfaction in the post-study interviews. *"The (AS-mode) MepsBot looks quite objective, and I trust it. With its help, I do not need to worry that I would say something wrong to those who could be vulnerable"* (AP6, F, age: 26). *"I will feel that my comment is reliable if (AS-mode) MepsBot gives me high scores"* (AP13, N/A, 22). *"Referring to the related content in (RE-mode) comments makes me more sure of my wordings"* (RP1, M, 24).

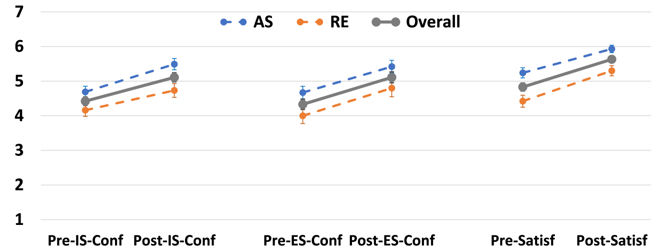

Figure 4. Means and standard errors of users' confidence in IS and ES and satisfaction with pre- and post-comments on a 7-point Likert scale (e.g., 1 not at all confident - 7 extremely confident).

Behaviors. On average, participants in the RE group spend more time on drafting one comment ($M = 10.0$mins, $SD = 4.3$mins) than those in the AS group ($M = 9.3$mins, $SD = 3.0$mins), but the difference is not significant (independent-samples t-test, $t_{88} = -.83, p = .41$). Across the 45 sessions (3 posts $\times$ 15 users) in each group, participants in the AS group click on "Back to edit" at least once in 17 sessions and indeed modify their comments in 16 sessions, while in the RE group there are 39 and 36, respectively. In the AS mode, the main reason for a direct submission without further revision is that MepsBot gives high scores on both IS and ES to the original comment, as reported by seven participants. *"It said that I wrote an excellent comment, but I could make it even better with words like 'you have'. I did not think that such modification was necessary as I felt great with high scores"* (AP3, M, 24). However, another five participants say that they would try to refine the comments according to MepsBot's hint, especially when the initial IS or ES score is not high. *"In the 3rd post, I got a medium ES score, and then I adjusted my comment to include the positive keywords it*

	AS: IS		AS: ES		RE: IS		RE: ES	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
High	42	43	36	42	40	42	30	37
Medium or Low	3	2	9	3	5	3	15	8

Table 3. IS & ES scores of pre- and post-comments judged by MepsBot.

suggested" (AP10, F, 24). In the RE mode, participants often update their comments as they *"find something useful in the suggested comments"* (RP1, M, 26). *"I realized that I might be too objective. It had good examples with cheerful words. I adopted them in my own one"* (RP10, F, 23).

Outcome of Comments (RQ1b)

Length and quality measured by MepsBot. The post-comments in AS group have slightly more words on average ($M = 117.9, SD = 42.5$) than pre-comments ($M = 113.6, SD = 45.8$). The difference is much larger in RE group, with an average of 104.0 ($SD = 44.5$) words in post-comments and 84.3 ($SD = 42.6$) words in pre-comments. According to MepsBot's support classifiers, most of the pre- and post-comments written by our participants get high IS and ES scores (Table 3). The numbers of high IS post-comments in AS and RE modes are rather close, but high ES post-comments in the AS group outnumbers those in the RE group by 42 to 37. There are six cases in which the ES scores are improved in the AS group and seven similar cases in the RE group. These results suggest that MepsBots could stimulate changes in the comments.

Changes in content. In total, we record 113 changed sentences, with 39 in the AS group and 74 in the RE group. The most common operation is adding a sentence ($N = 68$), followed by modifying ($N = 38$) and removing ($N = 7$) a sentence. As for what types of content are changed in these 113 sentences, three key themes emerge from thematic analysis: *IS-related*, *ES-related*, and *expression-related*. We count the occurrences of codes under different categories in each theme in AS and RE groups, as shown in Table 4. The changed content(s) in a sentence may be assigned to multiple codes regarding different aspects and thus fall into multiple categories. For instance, the modified sentence "Everyone (have → has) pressure and it (matters about → makes you stronger if you learn) how (you → to) deal with it" has one code (1st part) in category 5 and one code (2nd + 3rd parts) in category 6. The changes in content that can not be assigned to the categories in each theme are grouped into an extra N/A category.

1) IS-related Changes. Twenty (6 in AS, 14 in RE) participants refine the IS-related content with MepsBot's assistance. This happens more often in the RE group with 24 out of 74 changes on providing advice (32.4%) and 14 changes on sharing information (18.9%), compared to 5 (12.8%) and 7 (17.9%) out of 39 changes in the AS group. Eight participants mention that the suggested comments from the RE-mode MepsBot provide good insights into how to manifest IS in response to the posts. *"I found someone wrote a good point that constant dripping wears away a stone. I was inspired and suggested the buddy to record daily progress, however small"* (RP9, F, 24). The "social" words suggested by AS-mode MepsBot can also be inspiring, as mentioned by AP7 (M, 30), *"My mind was broadened by the suggested words about family and friends. I can share how people nearby work this problem out"*.

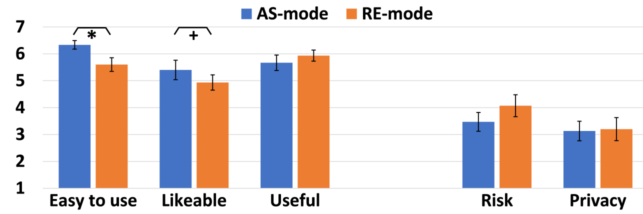

Figure 5. Means and standard errors of provider's perception of MepsBot in terms of acceptance and user experience (left), as well as concerns about potential risks and privacy (right) on a 7-point Likert scale (1 strongly disagree - 7 strongly agree); $*: p < .05, +: .05 < p < .1$.

2) ES-related Changes. Revisions also occur in ES-related contents after checking MepsBot's messages; more so in the RE group. There are 15 out of 74 changes (20.3%) about encouraging the help seekers and 13 changes (17.6%) about showing understanding in the RE group, while the numbers in AS group are 6 (15.4%) and 5 (12.8%) out of 39 changes. Seven participants point out that RE-mode MespBot can foster their manifestation of emotional support. *"I will check the emotional sentences with red highlighted words and consider using those sentences to cheer up the poster as I might be too rational"* (RP5, F, 21). Likewise, the "positive emotion" words appearing in AS-mode MepsBot's feedback could be helpful for boosting ES, as noted by AP8 (N/A, 25), *"The suggested words are positive, and I will try to use them in my comments to express my support"*.

3) Expression-related Changes. In addition to the IS- and ES-related content, MepsBot also prompts participants to correct typos (AS: 15.4%, RE: 10.8%) and polish wording in sentences (AS: 25.6%, RE: 12.2%), with higher frequencies in the AS group. This suggests that MepsBot (especially the AS-mode one) could make participants more reflective on their writing [70], and improve the readability and appropriateness of the submitted comment. *"The (AS-mode) MepsBot gave me a tip on the usage of 'if I were you', which sounds more appropriate, and I used it"* (AP0, M, 21). *"I paid more attention to the expression rather than the content of the (RE-mode MepsBot) recommended examples, and I would try to make my wording look natural"* (RP3, M, 28).

Perception towards MepsBot and Concerns (RQ2)

Acceptance and user experience. Figure 5 (left) shows the participants' ratings on their experience with MepsBot in two modes. In comparison to the RE-mode MepsBot ($M = 5.60, SD = 0.99$), the AS-mode one ($M = 6.33, SD = 0.62$) is rated significantly easier to use; independent-samples t-test, $t_{28} = 2.44, p < .05$; ***H2a*** accepted. The difference in the perceived likability of MepsBot is only marginally significant, with higher ratings for AS-mode ($M = 5.40, SD = 1.18$) than RE-mode ($M = 4.60, SD = 1.18$) MepsBots; $t_{28} = 1.85, p = .075$; ***H2b*** not accepted. Moreover, there is no significant difference between two modes (AS: $M = 5.67, SD = 1.11$; RE: $M = 5.93, SD = 0.80$) regarding perceived usefulness in helping users to write comments to support others online; $t_{28} = -.754, p = .457$; ***H2c*** not accepted. Participants really appreciate the simplicity of AS-mode MepsBot, saying that it *"is handy"* (AP10, F, 24) and *"only needs a click and a quick*

Theme	Category	Code example	**AS** (39 *sens*)	**RE** (74 *sens*)
IS-related	1. Advice	(Add)"Talk to seniors for help, who might have overcome the similar situation."; "[...] step by step ('-' $\rightarrow$, like reading more books in different areas)."	5 (12.8%)	24 (32.4%)
	2. Information	(Add)"Playing sports can exercise your body, more importantly, you can forget all of the pain in your brain."	7 (17.9%)	14 (18.9%)
ES-related	3. Encouragement	(Add)"Good luck! Stay positive."; (Add)"I am sure everything will go smoothly for you."	6 (15.4%)	15 (20.3%)
	4. Understanding	"I have similar situation as you ('-' $\rightarrow$, so do many of my friends):"; (Add)"Thanks for sharing your experience with us, I think I have the same problem."	5 (12.8%)	13 (17.6%)
Expression	5. Correct typo	"hourse" $\rightarrow$ "house"; "You have to lighten it up by yourself with (accumulation $\rightarrow$ accumulated) commitment to research, with passion, wish, and confidence."	6 (15.4%)	8 (10.8%)
	6. Polish sentence	"It is (amazing $\rightarrow$ impressive) and after so many (experiences $\rightarrow$ attempts)"; "you could also do ('-' $\rightarrow$ very) well in your research"	10 (25.6%)	9 (12.2%)
	7. N/A	[Test the system:] (Add)"If I were you I would rather xxx."	4 (10.3%)	2 (2.7%)

Table 4. Number of *sen*tences that are changed (i.e., add, remove or modify) during re-editions and types of changed contents in these *sen*tences, with frequencies of occurrence in each mode. Note that one *sen*tence may have different types of changed contents. '-' means blank.

glance" (AP12, M, 24). On the contrary, the RE-mode MepsBot sometimes *"could be confusing and give no direction"* (RP1, M, 26). Nevertheless, RE-mode MepsBot can be useful for helping providers write comments in various ways (some are mentioned in the previous subsections), one of which is that *"It feels like many people are discussing the issue together, and I enjoy it"* (RP15, M, 23).

Concerns - perceived accuracy. **1)** In the **AS-mode** group, MepsBot's accuracy on assessing the levels of support in the comments is rated favorably: 12 (13) of the 15 participants report its accuracy in predicting IS (ES) as 70% or higher, and the others rate it between 50%-70%. *"It is quite accurate. One time I got a medium ES score. I went back and found that I did write some words that could upset the support seeker"* (AP10, F, 24). Interestingly, AP6 (F, 26) tried to test the ability of MepsBot (Table 4 category 7) and found that *"it was robust"*. She added, *"I think it provides a standard for checking the support in the comment"*. Participants also provide possible justifications for cases when MepsBot is perceived inaccurate. AP2 (F, 23), who got high IS and ES scores in all three tasks, thought that *"it is unlikely that people get a high score every time. If that is the case, they will have less trust in the bot"*. The opaque mechanism of the classifiers also cause low accuracy ratings, as commented by AP5 (F, 25), *"I do not know how it assesses my scores, and I am not convinced"*.

For the 15 participants in the **RE-mode** group, six of them report MepsBot's accuracy in recommending similar comments to theirs as 70% or higher, five between 60%-70%, and four lower than 50%. Five participants mention that the suggested examples are somehow irrelevant to their comments, as RP8 (F, 20) explains, *"I feel like the recommendation is based on the similarity of words, but I can find little related content"*. This may be due to the two-step method used in RE-mode MepsBot, which only calculates semantic similarities for 50 potential candidates to enable real-time computation in a normal laptop. RP11 (M, 24) feels that *"the suggested comments are too long"*, and he only *"checked a few highlighted words and found them unmatched"* to his comment.

Concerns - perceived risks. As shown in Figure 5 (right), participants moderately worry that using MepsBot for writing support has potential risks, with an average rating $3.47(SD = 1.36)$ in the AS group and $4.07(SD = 1.59)$ in the RE group; no significant difference found between the two groups. **1)** For **AS-mode** MepsBot, four participants fear that **a)** the assessed dimensions are insufficient to reflect the subtle usage of language. *"It is better than none, but for the mental health issues, some words may look positive but could convey harmful meanings if misused"* (AP15, M, 23). AP12 (M, 24) also points out that *"people may add unnecessary contents purposely to get high scores from the bot"*. **b)** Regarding the potential frustration towards being judged [38, 30], none of our participants in the AS group mentioned that they were discouraged by MepsBot. This could be because they generally *"treat the commenting process very seriously"* (AP16, N/A, 25) and get high scores from MepsBot most of the time (Table 3).

2) For **RE-mode** MepsBot, the main risk lies in **a)** the potential misleading content in the suggested comments (eleven participants). *"The advice in the recommendations is not from professionals, which may have an adverse effect if conveyed to the seekers"* (RP1, M, 26). Eight participants also worry that **b)** there would be a lack of sincerity and authorship [32, 40] in the comments if RE-mode MepsBots are widely used in a community. *"I am afraid that it will make the community full of similar comments, which are insincere and invaluable for the support seekers"* (RP4, M, 25).

3) For **privacy issues** about MepsBot, participants rate average $3.13(SD = 1.41)$ for AS-mode and $3.20(SD = 1.66)$ for RE-mode MepsBots (Figure 5 right). *"I am not too concerned about that, since these posts and comments are public in the community. However, many people may worry that the company behind the bot would leak out their information"* (RP3, M, 28). *"I do not worry about the potential leak of my information. I do worry that my comments are stored in the bot's server forever and I cannot delete them"* (AP13, N/A, 22).

In summary, the support providers' confidence in and satisfaction with their comments are significantly improved after interacting with MepsBot. AS-mode MepsBot can stimulate providers to refine their expressions, while the RE-mode one encourages more re-editions on IS- and ES-related contents. Participants perceive AS- and RE-mode MepsBots differently and raise some concerns. These results indicate that technology-assisted manifestation of peer support in OMHCs is effective, but its mechanism needs to be carefully designed.

DISCUSSION

Design Considerations

From our experimental findings, we derive several design considerations for technological writing assistants in online mental health communities.

Offer assistance proactively during the commenting process

Three of the 30 participants noted that they often encountered difficult moments when they *"wrote a few sentences but could not finish the rest"* (AP13, N/A, 22). MepsBot could intervene more proactively rather than wait for the user's trigger in such a case, but it should first verify with users if they need help [57], like asking *"would you like to have a tip"* (RP4, M, 25).

Adjust assisting mechanisms based on the quality of comment

In our experiment, the AS-mode MepsBot boosted participants' confidence in their comments (Figure 4) especially when the reported IS and ES scores are high. The RE-mode MepsBot broadened users' mind to add contents to show IS and ES (Table 4). We suggest that MepsBot could adjust its assisting mechanisms according to the quality of the current draft. For example, MepsBot can act in AS mode if the comment is of high quality. When the comment is of low IS/ES or is too short, the MepsBot can activate its RE mechanism to offer some good examples for reference.

Customize quality metrics to user's personal needs

To improve the comments' quality, our MepsBot adopts the metrics of social support and promotes IS and ES through either AS or RE mode. However, there could be other dimensions for evaluating a comment, including but not limited to *"relevance to the post topic"* (RP9, F, 24) [23, 72], *"grammar"* (RP3, M, 28), similarity to others, comment's sentiment [58], and linguistic norms in the community [12]. Users may also want to prioritize different types of quality metrics according to their needs, e.g., *" I would like the MepsBot to emphasize IS and downplay ES as I just want to give some personal advice"* (AP12, M, 24). Therefore, we suggest that writing support technologies in OMHCs could offer a set of quality metrics for users to customize the direction in which they want to improve their comments. For example, the users who have poor writing skills can have MepsBot check their grammar first, and those who want to write a positive comment can call the bot to examine its sentiment.

Increase transparency of the technology

As presented in the RESULT section, participants' concern about system accuracy and privacy is mainly caused by the fact that they do not know how MepsBot works and how it uses their data. Also, participants in the RE group raise questions about the misleading information in the examples and the sincerity of their comments if following the recommendations. It is thus necessary for the writing support technology to be more transparent, communicating to users what and how well it can do as well as what are the consequences of their actions [2]. For example, when reporting the IS and ES score, MepsBot could tell the user that the classifiers are trained on comments labelled by experts. When recommending example comments, it could gently inform its limitation and remind the user to use the examples only for reference [50, 55].

Teach (psychotherapeutic) skills to show support

Ten participants (3 in AS, 7 in RE) reported that they learned how to manifest peer support in the comment with MepsBot's assistance. *"I learned to properly encourage the support seeker to try a therapy after I checked the example comments, which demonstrate a polite way to give out such an advice"* (RP5, F, 21). It suggests that the writing support tools like MepsBot were able to teach users peer support skills in situ. Such tools could further incorporate psychotherapeutic techniques (e.g., problem-solving therapy) into the suggestions [53] to nudge the users to write clinically effective comments.

Limitation and Future Work

Our work has several limitations. First, we only use data from one type of OMHC. Members in other mental heath communities may have different needs and expectations for real-time writing support. Second, we conduct our experiment only with students and restrict the post topics to those directly related to them. In a real OMHC, the demographics of members (e.g., pregnant women [33]) and the types of issues they are facing with are much more diverse. Third, most of our participants are new to the specific OMHC, while previous research suggests that newcomers and old-timers could have different commitment behaviors in OMHCs [6, 75]. In the future, we will explore the use of MepsBot in real-world OMHCs with different user populations. Fourth, though highlighting the most predictive features of IS and ES in the comments is perceived helpful, it is needed to have a systematic study to check the importance of these features in the future. Fifth, we do not include the condition in which the MepsBot only inserts a pause in the submission flow [71]. Future work needs to examine this condition to verify if a pause before submission is sufficient to improve the comments' quality.

CONCLUSION

To probe how support providers perceive and react to the on-the-fly writing support technologies in online mental health communities (OMHCs), we design MepsBot that helps providers write comments to show informational and emotional support. We examine two representative assisting mechanisms, i.e., assessment (AS) and recommendation (RE), in MepsBot through a mixed-design study. Results show that both MepsBots can boost users' confidence in and satisfaction with their drafted comments. AS-mode MepsBot is perceived as easier to use for its simplicity and can encourage users to refine expressions, while RE-mode MepsBot inspires providers to re-edit support-related contents. Our study demonstrates the value of writing support tools in OMHCs, and identifies users concerns about the accuracy and potential risks. We discuss ways to improve user experience and the effectiveness of future writing support technologies in OMHCs.

ACKNOWLEDGEMENT

We thank Taewook Kim and anonymous reviewers for the valuable feedback. This work is supported by the Research Grants Council of the Hong Kong Special Administrative Region, China under Grant No. 16204819.

REFERENCES

[1] 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. DOI: http://dx.doi.org/10.1162/tacl_a_00111

[2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 3, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300233

[3] Nazanin Andalibi and Andrea Forte. 2018. Responding to Sensitive Disclosures on Social Media: A Decision-Making Framework. *ACM Trans. Comput.-Hum. Interact.* 25, 6, Article 31 (Dec. 2018), 29 pages. DOI: http://dx.doi.org/10.1145/3241044

[4] Gerhard Andersson and Nickolai Titov. 2014. Advantages and limitations of Internet-based interventions for common mental disorders. *World psychiatry : official journal of the World Psychiatric Association (WPA)* 13 (02 2014), 4–11. DOI: http://dx.doi.org/10.1002/wps.20083

[5] Paul André, m.c. schraefel, Jaime Teevan, and Susan T. Dumais. 2009. Discovery is Never by Chance: Designing for (Un)Serendipity. In *Proceedings of the Seventh ACM Conference on Creativity and Cognition (C&C '09)*. ACM, New York, NY, USA, 305–314. DOI: http://dx.doi.org/10.1145/1640233.1640279

[6] Jaime Arguello, Brian S. Butler, Elisabeth Joyce, Robert Kraut, Kimberly S. Ling, Carolyn Rosé, and Xiaoqing Wang. 2006. Talk to Me: Foundations for Successful Individual-group Interactions in Online Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*. ACM, New York, NY, USA, 959–968. DOI: http://dx.doi.org/10.1145/1124772.1124916

[7] MHFA Australia. 2019. Mental Health First Aid. (2019). Retrieved June 1, 2019 from https://mhfa.com.au/.

[8] Shrey Bagroy, Ponnurangam Kumaraguru, and Munmun De Choudhury. 2017. A Social Media Based Index of Mental Well-Being in College Campuses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1634–1646. DOI: http://dx.doi.org/10.1145/3025453.3025909

[9] Piet Bracke, Wendy Christiaens, and Mieke Verhaeghe. 2008. Self-Esteem, Self-Efficacy, and the Balance of Peer Support Among Persons With Chronic Mental Health Problems. *Journal of Applied Social Psychology* 38, 2 (2008), 436–459. DOI: http://dx.doi.org/10.1111/j.1559-1816.2008.00312.x

[10] Michael Braun, Anja Mainz, Ronee Chadowitz, Bastian Pfleging, and Florian Alt. 2019. At Your Service: Designing Voice Assistant Personalities to Improve Automotive User Interfaces. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 40, 11 pages. DOI: http://dx.doi.org/10.1145/3290605.3300270

[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. DOI: http://dx.doi.org/10.1191/1478088706qp063oa

[12] Stevie Chancellor, Andrea Hu, and Munmun De Choudhury. 2018. Norms Matter: Contrasting Social Support Around Behavior Change in Online Weight Loss Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 666, 14 pages. DOI: http://dx.doi.org/10.1145/3173574.3174240

[13] Stevie Chancellor, Yannis Kalantidis, Jessica A. Pater, Munmun De Choudhury, and David A. Shamma. 2017. Multimodal Classification of Moderated Online Pro-Eating Disorder Content. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3213–3226. DOI: http://dx.doi.org/10.1145/3025453.3025985

[14] Stevie Chancellor, Zhiyuan (Jerry) Lin, and Munmun De Choudhury. 2016a. "This Post Will Just Get Taken Down": Characterizing Removed Pro-Eating Disorder Social Media Content. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1157–1162. DOI: http://dx.doi.org/10.1145/2858036.2858248

[15] Stevie Chancellor, Tanushree Mitra, and Munmun De Choudhury. 2016b. Recovery Amid Pro-Anorexia: Analysis of Recovery in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2111–2123. DOI: http://dx.doi.org/10.1145/2858036.2858246

[16] Stevie Chancellor, George Nitzburg, Andrea Hu, Francisco Zampieri, and Munmun De Choudhury. 2019. Discovering Alternative Treatments for Opioid Use Recovery Using Social Media. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 124, 15 pages. DOI: http://dx.doi.org/10.1145/3290605.3300354

[17] Stevie Chancellor, Jessica Annette Pater, Trustin Clear, Eric Gilbert, and Munmun De Choudhury. 2016. #Thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder Communities. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social*

Computing (CSCW ’16). ACM, New York, NY, USA, 1201–1213. DOI: http://dx.doi.org/10.1145/2818048.2819963

[18] Pamara Chang, Janis Whitlock, and Natalya Bazarova. 2018. “To Respond or not to Respond, that is the Question”: The Decision-Making Process of Providing Social Support to Distressed Posters on Facebook. *Social Media + Society* 4 (02 2018), 205630511875929. DOI:http://dx.doi.org/10.1177/2056305118759290

[19] Munmun De Choudhury and Emre Kiciman. 2017. The Language of Social Support in Social Media and Its Effect on Suicidal Ideation Risk. *Proceedings of the ... International AAAI Conference on Weblogs and Social Media. International AAAI Conference on Weblogs and Social Media* 2017 (2017), 32–41.

[20] Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A. Smith. 2018. Creative Writing with a Machine in the Loop: Case Studies on Slogans and Stories. In *23rd International Conference on Intelligent User Interfaces (IUI ’18)*. ACM, New York, NY, USA, 329–340. DOI: http://dx.doi.org/10.1145/3172944.3172983

[21] Munmun De Choudhury, Scott Counts, Eric J. Horvitz, and Aaron Hoff. 2014. Characterizing and Predicting Postpartum Depression from Shared Facebook Data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW ’14)*. ACM, New York, NY, USA, 626–638. DOI:http://dx.doi.org/10.1145/2531602.2531675

[22] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics, Minneapolis, Minnesota, 4171–4186. DOI: http://dx.doi.org/10.18653/v1/N19-1423

[23] Karthik Dinakar, Jackie Chen, Henry Lieberman, Rosalind Picard, and Robert Filbin. 2015. Mixed-Initiative Real-Time Topic Modeling & Visualization for Crisis Counseling. In *Proceedings of the 20th International Conference on Intelligent User Interfaces (IUI ’15)*. ACM, New York, NY, USA, 417–426. DOI: http://dx.doi.org/10.1145/2678025.2701395

[24] Keith S. Dobson. 2010. *Handbook of cognitive-behavioral therapies, 3rd ed.* Guilford Press, New York, NY, US. xiv, 481–xiv, 481 pages.

[25] 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 (CHI ’12)*. ACM, New York, NY, USA, 1421–1430. DOI: http://dx.doi.org/10.1145/2207676.2208602

[26] Bryan Dosono and Bryan Semaan. 2019. Moderation Practices As Emotional Labor in Sustaining Online Communities: The Case of AAPI Identity Work on Reddit. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI ’19)*. ACM, New York, NY, USA, Article 142, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300372

[27] Elasticsearch. 2019. Python Elasticsearch Client – Elasticsearch 7.9.9 documentation. (2019). Retrieved May 23, 2019 from https://mhfa.com.au/.

[28] John F M Gleeson, Mario Alvarez-Jimenez, and Reeva Lederman. 2012. Moderated Online Social Therapy for Recovery From Early Psychosis. *Psychiatric services (Washington, D.C.)* 63 (07 2012), 719. DOI: http://dx.doi.org/10.1176/appi.ps.20120p719

[29] Noura Farra, Swapna Somasundaran, and Jill Burstein. 2015. Scoring Persuasive Essays Using Opinions and their Targets. In *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications*. Association for Computational Linguistics, Denver, Colorado, 64–74. DOI: http://dx.doi.org/10.3115/v1/W15-0608

[30] Jessica L. Feuston and Anne Marie Piper. 2019. Everyday Experiences: Small Stories and Mental Illness on Instagram. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI ’19)*. ACM, New York, NY, USA, Article 265, 14 pages. DOI:http://dx.doi.org/10.1145/3290605.3300495

[31] Anna Geraedts, Annet Kleiboer, Jos Twisk, Noortje Wiezer, Willem van Mechelen, and Pim Cuijpers. 2014. Long-Term Results of a Web-Based Guided Self-Help Intervention for Employees With Depressive Symptoms: Randomized Controlled Trial. *Journal of Medical Internet Research* (07 2014). DOI: http://dx.doi.org/10.2196/jmir.3539

[32] Katy Ilonka Gero and Lydia B. Chilton. 2019. Metaphoria: An Algorithmic Companion for Metaphor Creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI ’19)*. ACM, New York, NY, USA, Article 296, 12 pages. DOI: http://dx.doi.org/10.1145/3290605.3300526

[33] Xinning Gui, Yu Chen, Yubo Kou, Katie Pine, and Yunan Chen. 2017. Investigating Support Seeking from Peers for Pregnancy in Online Health Communities. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 50 (Dec. 2017), 19 pages. DOI: http://dx.doi.org/10.1145/3134685

[34] Tayebeh Fasihi Harandi, Maryam Mohammad Taghinasab, and Tayebeh Dehghan Nayeri. 2017. The correlation of social support with mental health: A meta-analysis. In *Electronic physician*.

[35] Thomas K. Houston, Lisa A Cooper, and Daniel Ernest Ford. 2002. Internet support groups for depression: a 1-year prospective cohort study. *The American journal of psychiatry* 159 12 (2002), 2062–8.

[36] Julie S. Hui, Darren Gergle, and Elizabeth M. Gerber. 2018. IntroAssist: A Tool to Support Writing Introductory Help Requests. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 22, 13 pages. DOI: http://dx.doi.org/10.1145/3173574.3173596

[37] Lisa J Barney, Kathleen Griffiths, and Michelle Banfield. 2011. Explicit and implicit information needs of people with depression: A qualitative investigation of problems reported on an online depression support forum. *BMC psychiatry* 11 (05 2011), 88. DOI: http://dx.doi.org/10.1186/1471-244X-11-88

[38] Ryan Kelly, Daniel Gooch, and Leon Watts. 2018. 'It's More Like a Letter': An Exploration of Mediated Conversational Effort in Message Builder. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 87 (Nov. 2018), 23 pages. DOI: http://dx.doi.org/10.1145/3274356

[39] Jan Kietzmann, Kristopher Hermkens, Ian McCarthy, and Bruno Silvestre. 2011. Social Media? Get Serious! Understanding the Functional Building Blocks of Social Media. *Business Horizons* 54 (05 2011), 241–251. DOI: http://dx.doi.org/10.1016/j.bushor.2011.01.005

[40] Taewook Kim, Jungsoo Lee, Zhenhui Peng, and Xiaojuan Ma. 2019. Love in Lyrics: An Exploration of Supporting Textual Manifestation of Affection in Social Messaging. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 79 (Nov. 2019), 27 pages. DOI: http://dx.doi.org/10.1145/3359181

[41] Joseph Lampel and Ajay Bhalla. 2007. The Role of Status Seeking in Online Communities: Giving the Gift of Experience. *Journal of Computer-Mediated Communication* 12, 2 (2007).

[42] Reeva Lederman, Greg Wadley, John Gleeson, Sarah Bendall, and Mario Alvarez-Jimenez. 2014. Moderated Online Social Therapy: Designing and Evaluating Technology for Mental Health. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21 (02 2014). DOI:http://dx.doi.org/10.1145/2513179

[43] Minha Lee, Sander Ackermans, Nena van As, Hanwen Chang, Enzo Lucas, and Wijnand IJsselsteijn. 2019. Caring for Vincent: A Chatbot for Self-Compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 702, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300932

[44] Guo Li, Xiaomu Zhou, Tun Lu, Jiang Yang, and Ning Gu. 2016. SunForum: Understanding Depression in a Chinese Online Community. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 515–526. DOI: http://dx.doi.org/10.1145/2818048.2819994

[45] Bernd Loewe, K Urt Kroenke, Wolfgang Herzog, and Kerstin Graefe. 2004. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). *Journal of affective disorders* 81 1 (2004), 61–6.

[46] Ann Macaskill. 2013. The mental health of university students in the United Kingdom. *British Journal of Guidance & Counselling* 41, 4 (2013), 426–441. DOI: http://dx.doi.org/10.1080/03069885.2012.743110

[47] Shery Mead, David Hilton, and Laurie Curtis. 2001. Peer Support: A Theoretical Perspective. *Psychiatric rehabilitation journal* 25 (02 2001), 134–41. DOI: http://dx.doi.org/10.1037/h0095032

[48] Anne Moorhead, Diane Hazlett, Laura Harrison, Jennifer K Carroll, Anthea Irwin, and Ciska Hoving. 2013. A New Dimension of Health Care: Systematic Review of the Uses, Benefits, and Limitations of Social Media for Health Communication. *Journal of medical Internet research* 15 (04 2013), e85. DOI: http://dx.doi.org/10.2196/jmir.1933

[49] Robert Morris, Stephen Schueller, and Rosalind Picard. 2015. Efficacy of a Web-Based, Crowdsourced Peer-To-Peer Cognitive Reappraisal Platform for Depression: Randomized Controlled Trial. *Journal of Medical Internet Research* 17 (03 2015), e72. DOI: http://dx.doi.org/10.2196/jmir.4167

[50] Robert R. Morris, Kareem Kouddous, Rohan Kshirsagar, and Stephen Matthew Schueller. 2018. Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions. In *Journal of medical Internet research*.

[51] John Naslund, Kelly Aschbrenner, LA Marsch, and Stephen Bartels. 2016. The future of mental health care: Peer-to-peer support and social media. *Epidemiology and Psychiatric Sciences* -1 (01 2016), 1–10. DOI: http://dx.doi.org/10.1017/S2045796015001067

[52] Kathleen O'Leary, Arpita Bhattacharya, Sean A. Munson, Jacob O. Wobbrock, and Wanda Pratt. 2017. Design Opportunities for Mental Health Peer Support Technologies. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM, New York, NY, USA, 1470–1484. DOI: http://dx.doi.org/10.1145/2998181.2998349

[53] Kathleen O'Leary, Stephen M. Schueller, Jacob O. Wobbrock, and Wanda Pratt. 2018. "Suddenly, We Got to Become Therapists for Each Other": Designing Peer Support Chats for Mental Health. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 331, 14 pages. DOI: http://dx.doi.org/10.1145/3173574.3173905

[54] World Health Organization. 2017. Mental health: massive scale-up of resources needed if global targets are to be met. (2017). Retrieved June 12, 2019 from https://www.who.int/mental_health/evidence/atlas/atlas_2017_web_note/en/.

[55] Steffi Paepcke and Leila Takayama. 2010. Judging a bot by its cover: An experiment on expectation setting for personal robots. *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2010), 45–52.

[56] Sung Youl Park. 2009. An Analysis of the Technology Acceptance Model in Understanding University Students' Behavioral Intention to Use e-Learning. *Educational Technology & Society* 12 (07 2009), 150–162.

[57] Zhenhui Peng, Yunhwan Kwon, Jiaan Lu, Ziming Wu, and Xiaojuan Ma. 2019. Design and Evaluation of Service Robot's Proactivity in Decision-Making Support Process. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 98, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300328

[58] 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 (CHI '19)*. ACM, New York, NY, USA, Article 64, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300294

[59] Pushshift. 2019. Reddit Statistics - pushshift.io. (2019). Retrieved June 1, 2019 from https://pushshift.io/.

[60] Yla R. Tausczik and James Pennebaker. 2010. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology* 29 (03 2010), 24–54. DOI:http://dx.doi.org/10.1177/0261927X09351676

[61] Reddit. 2019a. Reddiquette: Please do and please don't. (2019). Retrieved September 2nd, 2019 from https://www.reddit.com/wiki/reddiquette/.

[62] Reddit. 2019b. *r/depression*, a supportive space for anyone struggling with depression. (2019). Retrieved March 25, 2019 from https://www.reddit.com/r/depression/.

[63] 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 (CHI '19)*. ACM, New York, NY, USA, Article 425, 14 pages. DOI: http://dx.doi.org/10.1145/3290605.3300655

[64] Sherry Ruan, Liwei Jiang, Justin Xu, Bryce Joe-Kun Tham, Zhengneng Qiu, Yeshuang Zhu, Elizabeth L. Murnane, Emma Brunskill, and James A. Landay. 2019. QuizBot: A Dialogue-based Adaptive Learning System for Factual Knowledge. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 357, 13 pages. DOI:http://dx.doi.org/10.1145/3290605.3300587

[65] Mark Salzer, Nicole Darr, Gina Calhoun, William Boyer, Randall Loss, Jerry Goessel, Edward Schwenk, and Eugene Brusilovskiy. 2013. Benefits of working as a certified peer specialist: Results from a statewide survey. *Psychiatric rehabilitation journal* 36 (09 2013), 219–21. DOI:http://dx.doi.org/10.1037/prj0000016

[66] Shruti Sannon, Elizabeth L. Murnane, Natalya N. Bazarova, and Geri Gay. 2019. "I Was Really, Really Nervous Posting It": Communicating About Invisible Chronic Illnesses Across Social Media Platforms. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 353, 13 pages. DOI: http://dx.doi.org/10.1145/3290605.3300583

[67] Catherine E. Schaefer, James Charles Coyne, and Richard Stanley Lazarus. 1981. The health-related functions of social support. *Journal of Behavioral Medicine* 4 (1981), 381–406.

[68] 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 (CHI '18)*. ACM, New York, NY, USA, Article 641, 13 pages. DOI: http://dx.doi.org/10.1145/3173574.3174215

[69] Stefan Siersdorfer, Sergiu Chelaru, Wolfgang Nejdl, and Jose San Pedro. 2010. How Useful Are Your Comments?: Analyzing and Predicting Youtube Comments and Comment Ratings. In *Proceedings of the 19th International Conference on World Wide Web (WWW '10)*. ACM, New York, NY, USA, 891–900. DOI: http://dx.doi.org/10.1145/1772690.1772781

[70] Liuping Wang, Xiangmin Fan, Feng Tian, Lingjia Deng, Shuai Ma, Jun Huang, and Hongan Wang. 2018. mirrorU: Scaffolding Emotional Reflection via In-Situ Assessment and Interactive Feedback. In *CHI Extended Abstracts*.

[71] Yang Wang, Pedro Giovanni Leon, Alessandro Acquisti, Lorrie Faith Cranor, Alain Forget, and Norman Sadeh. 2014. A Field Trial of Privacy Nudges for Facebook. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2367–2376. DOI: http://dx.doi.org/10.1145/2556288.2557413

[72] Yi-Chia Wang, Robert Kraut, and John M. Levine. 2012. To Stay or Leave?: The Relationship of Emotional and Informational Support to Commitment in Online Health Support Groups. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*. ACM, New York, NY, USA, 833–842. DOI:http://dx.doi.org/10.1145/2145204.2145329

[73] Jen-Her Wu and Shu-Ching Wang. 2005. What Drives Mobile Commerce? *Inf. Manage.* 42, 5 (July 2005), 719–729. DOI: http://dx.doi.org/10.1016/j.im.2004.07.001

[74] Shaomei Wu, Lindsay Reynolds, Xian Li, and Francisco Guzmán. 2019. Design and Evaluation of a Social Media Writing Support Tool for People with Dyslexia. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 516, 14 pages. DOI: http://dx.doi.org/10.1145/3290605.3300746

[75] Diyi Yang, Robert Kraut, and John M. Levine. 2017. Commitment of Newcomers and Old-timers to Online Health Support Communities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6363–6375. DOI: http://dx.doi.org/10.1145/3025453.3026008

[76] Diyi Yang, Robert E. Kraut, Tenbroeck Smith, Elijah Mayfield, and Dan Jurafsky. 2019a. Seekers, Providers, Welcomers, and Storytellers: Modeling Social Roles in Online Health Communities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 344, 14 pages. DOI: http://dx.doi.org/10.1145/3290605.3300574

[77] Diyi Yang, Zheng Yao, Joseph Seering, and Robert Kraut. 2019b. The Channel Matters: Self-disclosure, Reciprocity and Social Support in Online Cancer Support Groups. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 31, 15 pages. DOI:http://dx.doi.org/10.1145/3290605.3300261