

Original Paper

Prevalence of Cognitive Distortion Markers in a Suicide Prevention Chat Service: Mixed Methods Study

Marijn ten Thij^{1,2}, PhD; Saskia Mérelle³, PhD; Renske Gilissen^{3,4}, Prof Dr; Johan Bollen^{2,5}, Prof Dr

¹Department of Intelligent Systems, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands

²Centre for Social and Biomedical Complexity, Indiana University Bloomington, Bloomington, IN, United States

³113 Suicide Prevention, Amsterdam, The Netherlands

⁴Instituut Psychologie, Faculteit der Sociale Wetenschappen, Leiden University, Leiden, The Netherlands

⁵Instituut voor Informatica, Faculteit Natuurwetenschappen, Wiskunde, en Informatica, University of Amsterdam, Amsterdam, The Netherlands

Corresponding Author:

Marijn ten Thij, PhD
Department of Intelligent Systems
Tilburg School of Humanities and Digital Sciences, Tilburg University
Warandelaan 2
Tilburg 5037 AB
The Netherlands
Email: m.c.tenthij@tilburguniversity.edu

Abstract

Background: Suicide helplines increasingly employ chat services to aid those in urgent need, but the wording and structure of text-driven exchanges may affect their effectiveness.

Objective: Given the association of cognitive distortions with depression and anxiety, this study investigated their prevalence in the language of individuals seeking help from the Dutch 113 suicide helpline.

Methods: We observed the prevalence of cognitive distortions for both help seekers and counselors in a large volume of chat sessions (N=71,148) of the Dutch 113 suicide chat helpline using natural language processing. The results were compared to 2 large collections of online text data from Dutch social media and web content.

Results: We found that nearly all types of cognitive distortions are more prevalent in the language of help seekers compared to the control group of helpline counselors. Distortions of the personalizing, emotional reasoning, and mental filtering types were, respectively, 20.22, 7.87, and 4.53 times more prevalent among help seekers, revealing a distinct pattern of thought and language among individuals affected by suicidality.

Conclusions: Our results raise the prospect of improving the effectiveness of online therapeutic interventions that target cognitive distortions through lexical analysis that detects the cognitive and lexical markers of suicidality.

JMIR Ment Health 2026;13:e81213; doi: [10.2196/81213](https://doi.org/10.2196/81213)

Keywords: cognitive distortions; online suicide prevention helpline; natural language processing; cognitive behavioral therapy; suicide prevention

Introduction

The Societal Burden of Suicide

Suicide exacts a staggering societal cost. Recent World Health Organization (WHO) estimates indicate that more than 700,000 individuals die of suicide each year and its worldwide mortality exceeds that of malaria, HIV/AIDS, war, and homicide [1]. Suicide is the 10th leading cause of death in the United States, where its economic costs are estimated to be over US \$70 billion per year in lifetime medical expenses

and loss of productivity alone, with similar numbers across the globe [2-7]. Moreover, suicide is a leading cause of death among young people in the Netherlands. Increasing suicide rates among young people have been observed [8].

Beyond deaths from suicide, many more individuals experience suicidal thoughts and intent. US Centers for Disease Control and Prevention (CDC) statistics indicate that, in 2018, a total of 10.7 million, 3.3 million, and 1.4 million American adults either seriously thought about suicide, made a plan, and/or attempted suicide, respectively [9]. The burden on minors and young adults is equally high. In 2019, a total of

18.8% of high school students reported seriously considering suicide and 8.9% reported an attempt [10].

Analyzing Language From Chat-Based Suicide Prevention Helplines

Given these indicators of a global mental health crisis [11], the death toll of suicide, and the large population contemplating or attempting suicide, scalable and effective mitigation is of the utmost importance [12-14]. Here, we focus on suicide prevention helplines, which aim to provide immediate assistance and access to important resources [15] to those experiencing an imminent risk of suicide. Many telephone helplines have recently been augmented with online chat services to extend access, in particular for the young [16] who may prefer chat services over telephone conversations. Given their use of ubiquitous online chat applications, chat helplines can offer time-sensitive, real-time, cost-effective, and anonymous assistance to individuals reaching out at a time of urgent need [17]. They are for these reasons preferred over other channels, such as email, text messaging, voice calls, and face-to-face counseling [18]. A large-scale study of a national suicide prevention helpline showed that the chat functionality had been helpful for almost 70% of the users and that 45% felt less suicidal after the interaction(s) [19]. In addition, short-term improvements were recently observed in suicidal ideation and psychological risk factors after a chat conversation in another national helpline [20].

Thus far, research has focused on studying suicidal ideation from language, for example, through linguistic features from social media content [21-24], poetry [25], forums [23,24,26-28], or suicide notes [29-31] using tools, such as Linguistic Inquiry and Word Count [32]. Examples of such linguistic features that are extracted from these texts are self-to-relational word counts [33], linguistic accommodation [26], hopelessness [28], social disengagement [28], or categories of stress [31].

By definition, chat-based suicide helplines rely on text-driven communication, which, although it takes place in real time, does not include verbal cues, body language, intonation, and other important aspects of in-person communication. They require that respondents interpret the written language of help seekers, detect cues and indications of the help seeker's state, and respond accordingly. However, it is not clear what these written language cues or features may be in the context of chat-driven suicide helplines; this remains an important knowledge gap. Recent efforts that have started to address this gap focus on providing suggestions for counselors based on previous chat sessions [34] or training counselors in motivational interviewing [35].

Detecting Cognitive Distortions in Suicide Prevention Helplines

There exists an extensive literature on language features associated with depression and other internalizing disorders, which are known to be strongly comorbid with suicidal ideation and depression [36-38]. For example, Eichstaedt et al [39] showed that language features indicative of emotional distress, interpersonal dissatisfaction, and

rumination can be leveraged by machine-learning algorithms to predict depression in Facebook users at a reasonable level of accuracy. Rude et al [40] showed that the language of individuals with depression is indeed marked by the prevalence of specific identifiable terms. However, the language of individuals with depression is not only marked by the presence of discrete features that may be useful in diagnostic classification tasks, but also by a style of thinking that is deeply implicated with the development of internalizing disorders such as depression.

According to the theory underlying cognitive behavioral theory, the cognitive triad and the cognitive mediation model [41-43], individuals with depression exhibit self-reinforcing thought patterns in which unrealistic and overly negative views about the world, the future, and oneself cause a cascade of detrimental cognitive, affective, and behavioral changes associated with internalizing disorders. These thought patterns are referred to as cognitive distortions and frequently described in terms of 12 general types, eg, "I will never amount to anything," which in this case amounts to "fortune-telling," "dichotomization," and "personalizing." Regardless of type, all cognitive distortions share the characteristic that they amount to unrealistic, exaggerated, and overly rigid thinking patterns.

Cognitive behavioral therapy (CBT), now the leading treatment of depression, anxiety, and suicide [44-49], considers cognitive distortions an important intervention target. In fact, the very effectiveness of CBT demonstrates the importance of the identification and mitigation of cognitive distortions in an individual seeking help with depression or anxiety. Previous research [50,51] demonstrated that the online language of individuals with depression is indeed marked by a higher prevalence of cognitive distortions than a random control sample. Since depression is strongly comorbid with suicide [36-38], here we investigate whether the written language of individuals seeking help with suicidal thoughts in a chat-based helpline is similarly impacted by the presence of cognitive distortions. If so, as indicated by the effectiveness of CBT [48,49], addressing cognitive distortions may be an important driver and intervention target for suicidality in this context.

We analyzed 6,349,202 chat messages recorded from July 2017 to June 2021, which were split between those written by either help seekers or the professional respondents of the 113 suicide helpline in the Netherlands. We found that markers of some types of cognitive distortions are 20-fold more prevalent in the language of help seekers than the 113 counselors. In fact, a comparison with the online language of individuals with depression and a random sample recorded on Twitter revealed a similar difference and indicates that Twitter language may be particularly enriched with cognitive distortions even in the absence of any diagnosed mental health issues.

Methods

Ethical Considerations

This study was approved by the Indiana University Bloomington Institutional Review Board under protocol 2007591998 “Retrospective language analysis of suicide prevention online chat content (Dutch 113 service),” which was reviewed specifically for this study and its research team. Help seekers did not receive any compensation for the inclusion of their sessions in our analysis. By participating in the chat sessions at 113, help seekers approve that these sessions can be used for research purposes, as stated in the terms and conditions.

As our analysis pertains to individual health-related information, we took additional steps to ensure the privacy of all individuals involved in the chat sessions. We deidentified all data by assigning each chat session a unique but anonymous label, for example, c1123456, to remove all identifying information. All times are reported in terms of time elapsed from the start of the conversation. All raw data are stored in a protected environment at 113 suicide prevention. Finally, all analyses were performed and reported in the aggregate, that is, without reference to the individuals that constitute the specific cohorts of the study.

Data and Sample

We performed our analysis on chat sessions recorded by the secure chat platform of the 113 suicide prevention helpline in the Netherlands. On this platform, help seekers can reach out to 113 employees to discuss any topics related to suicide. For our analysis, we selected all chat sessions that passed the initial triage stage determining whether the help seeker was in a safe location before being passed to a professional counselor (a psychologist or a volunteer). Chat sessions have 3 types of participants: help seekers, 113 counselors, and the chat service system itself. The chat service system posts automated messages to indicate that it has taken a specific action; we exclude these messages from our analysis. Applying this criterion resulted in a large-scale dataset consisting of 71,148 chat sessions that were recorded between July 1, 2017, and June 1, 2021. In total, these chat sessions contain 3,316,873 messages posted by help seekers and 3,032,329 messages posted by 113 counselors. The average monthly number of forwarded chat sessions that we have included in our analysis is 1513 (95% CI 1199-1874), showing high levels of engagement with the 113 helpline.

This study investigated the differences between the written language of a cohort of 113 help seekers who are seeking urgent help with suicidal ideation or intent and the written language of a control cohort with little or no suicidality in the context of an online chat environment. Given the nature of the 113 helpline, we did not expect to find many help seekers with low levels of suicidality that could serve as a control cohort. In fact, an analysis of the responses to the questionnaire [35,52] that help seekers are required to fill in showed insufficient discrimination between low and high levels of suicidality to construct a control cohort (see

section “Types of Help seekers as control” in [Multimedia Appendix 1](#)). Here, we chose to use the 113 counselors as a control group since they are expected to be least affected by suicidality and depression, while conversing with help seekers on the same platform about the same topic. For the remainder of the analysis, h denotes the help seeker cohort and O denotes the Counselor cohort.

In addition, we compared the written content of the 113 chat sessions with 2 very large samples of Dutch texts, the latter providing control data that are independent from a suicide prevention helpline. First, we included the CC100-Dutch dataset, a large collection of Dutch sentences (N=229,914,131), which were extracted from websites that were included in the January to December 2018 CommonCrawl snapshots from the CC-Net repository [53]. Second, we included a comparison to the language of a large sample of users on Twitter (now X), a popular microblogging platform. Using an open-access repository of COVID-19 messages [54], cohorts of active Twitter users in Dutch municipalities with more than 100,000 inhabitants in 2020 [55] were constructed based on the user-provided location information on their own profile page. Concretely, users were included who indicated that they live in one of the following cities: Amsterdam, Rotterdam, The Hague, Utrecht, Eindhoven, Groningen, Tilburg, Almere, Breda, Nijmegen, Apeldoorn, Haarlem, Arnhem, Hoofddorp, Amersfoort, Enschede, Zaanstad, Den Bosch, Zwolle, Leiden, Leeuwarden, Zoetermeer, Maastricht, Ede, Dordrecht, Alphen aan den Rijn, Alkmaar, Delft, Emmen, Venlo, and Deventer. To ensure that these Twitter users were indeed individuals, we used the M3 system [56] (a deep learning classifier trained on large collections of Twitter profiles) to classify accounts according to 3 categories: gender, age, and organization. Using the outcomes of the organization classifier, we only kept accounts that had a likelihood score of more than 0.8 of being an individual in our analysis. In total, this dataset consists of 4829 Twitter users for which the individual timelines (3200 most recent tweets, time-sorted by date posted by individual) were collected using Twitter’s application programming interface for an analytical sample of 4,600,054 tweets that were posted between July 1, 2017, and October 31, 2020. Using 0.8 as a threshold for gender and 0.6 as a threshold for age (as this category contains 4 labels rather than 2), we can assess the demographics of the Twitter users we have included in our analysis. In total, our sample contains 3119 male individuals (64 of whom were aged 18 years and younger, 117 aged 19-29 years, 455 aged 30-39 years, and 1959 aged 40 years and older) and 1522 female individuals (39 of whom were aged 18 years and younger, 145 aged 19-29 years, 202 aged 30-39 years, and 839 aged 40 years and older). Note that the totals do not necessarily have to sum up to the total amount of included Twitter users, as the outcomes of the M3 system are below the used threshold (0.8 for gender and 0.6 for age).

Cognitive Distortions and CBT

Cognitive distortions are maladaptive patterns of thought that individuals with depression exhibit in which unrealistic and overly negative views about the world, the future, and oneself

cause a cascade of detrimental cognitive, affective, and behavioral changes associated with internalizing disorders [41-43]. These cognitive distortions are frequently described in terms of 12 general types, for example, “I will never amount to anything,” which in this case amounts to “fortune-telling,” “dichotomization,” and “personalizing.” Regardless of type, all cognitive distortions share the characteristic that they amount to unrealistic, exaggerated, and overly rigid thinking patterns.

CBT, now the leading treatment of depression, anxiety, and suicide [44-49], considers cognitive distortions an important intervention target. In fact, the very effectiveness of CBT demonstrates the importance of the identification and mitigation of cognitive distortions in an individual seeking help with depression or anxiety.

Construction of Set of Cognitive Distortion Schemata N-Grams

The detection of cognitive distortions in written language, in open-form online chat sessions, remains an open question. Some recent advances suggest natural language processing and machine learning can be leveraged to detect the presence of cognitive distortions in text [57,58] at reasonable levels of accuracy, even matching the performance of human raters, but it is not clear how well these methods generalize to online chat data, nor whether the lexical features involved in the classification correspond to our understanding of

cognitive distortions [59]. In fact, some research indicates that sophisticated machine-learning models, such as deep learning, may in fact be outperformed by more transparent models that capture a small number of specific expressions associated with cognitive distortions [60].

Here, we used a theory-driven method proposed by Bathina et al [50], namely a lexicon of 241 one- to five-grams that was specifically designed and vetted by a panel of experts in CBT, originally introduced by Beck [41-43], to capture the sequences of words involved with the expression of 12 common types of cognitive distortions (see Table 1). For example, the 1-gram “never,” among other 1-grams, is deemed to be indicative of “dichotomization,” and the 3-grams “they will think” or “I am a” are deemed to be indicative of “mindreading” and “labeling,” respectively. A message that contains either one of the n-grams in the cognitive distortion schemata (CDS) lexicon is deemed to contain an indication of the expression of the specific cognitive distortion type, regardless of its topic or context. For example, the message “I will never be loved” contains the 3-gram “I will never,” which is associated with the expression of a cognitive distortion of the “fortune telling” type. Counting the number of matches of the entire CDS lexicon against the content of the chat-session messages (by help seekers or counselors) provides an estimate of the prevalence of cognitive distortion expressions in their language.

Table 1. Descriptive statistics of our set of cognitive distortion schemata (CDS), which are grouped in 12 common types (“distortion type”)^a.

Distortion type	Notation	N_{CDS}	N_{\exists}	N_{regex}
Catastrophizing	<i>c</i>	25	16	15
Dichotomous reasoning	<i>dr</i>	25	24	4
Disqualifying the positive	<i>ntp</i>	14	13	8
Emotional reasoning	<i>er</i>	10	10	3
Fortune-telling	<i>f</i>	8	5	8
Labeling and mislabeling	<i>lam</i>	47	45	35
Magnification and minimization	<i>mam</i>	9	9	2
Mental filtering	<i>mf</i>	16	15	10
Mindreading	<i>m</i>	72	63	63
Overgeneralizing	<i>o</i>	20	20	12
Personalizing	<i>p</i>	15	15	7
Should statements	<i>ss</i>	4	4	4
Total	<i>C</i>	265	239	171

^aThe column “ N_{CDS} ” indicates the number of schemata in the specific category, the column “ N_{\exists} ” shows the number of n-grams in each category that were actually found in the chat session data, and the column “ N_{regex} ” indicates the number of schemata that contain a regular expression.

The CDS lexicon is grouped in 12 distortion types following a common taxonomy, but its use does not rely on claims with respect to the component structure of cognitive distortions in general. Each n-gram is matched independently. The number of literal matches of the n-grams of the CDS lexicon in the messages of either help seekers or 113 counselors is normalized by message volume to determine CDS prevalence (see section “Measuring CDS in chat messages”).

The original CDS n-grams were translated from English to Dutch by 4 bilingual native language speakers of

Dutch, namely 2 Dutch-speaking coauthors, and in-house 113 helpline psychologists Wilco Janssen and Jeroen Gomes, to match the language in which the 113 helpline conversations were conducted. The resulting set of 4 translations (for each CDS) was combined to a single list of translated CDS n-grams through a process of iterative design. Finally, 2 Dutch experts in CBT (Claudi Bockting and Marcus Huibers) were asked to validate whether the translations still captured the distorted thinking.

The translations focused on retaining the distorted thinking aspect of the schemata, respecting grammar (eg, gender) and idiomatic expressions. For that reason, the number of Dutch schemata ($N=265$) does not equal the number of English schemata ($N=241$), since one-on-one translation was not always feasible. Moreover, due to differences in Dutch word

ordering (eg, verb at the end of the sentence), many of the schemata had to be formulated as regular expressions instead of n-grams. The number of schemata per category in the Dutch CDS set along with the number of observed schemata per category can be found in Table 1. The complete set of CDS is freely available in Table 2.

Table 2. Cognitive distortion types and the corresponding schemata in Dutch.

Distortion type	Cognitive distortion schemata
Catastrophizing	(zallgaat)(.+ l)falen, zal(+ l)fout gaan, zal(+ l)verkeerd gaan, zal(+ l)mis gaan, (zallgaat)(.+ l)mislukken, (zallgaat)(.+ l)eindigen, (zallgaat)(.+ l)stoppen, (zallgaat)(.+ l)ophouden, wordt onmogelijk, zal onmogelijk zijn, (zallgaat) niet (gebeuren)plaatsvinden, wordt vreselijk, zal vreselijk zijn, wordt verschrikkelijk, zal verschrikkelijk zijn, wordt een ramp, zal(+ l)een ramp zijn, (zallgaat) nooit (eindigen)stoppen, (stopt)leindigt niet, houdt niet op, gaat niet (stoppen)ophouden
Dichotomous reasoning	alleen, elke, iedere, iedereen, allemaal, alles, overal, altijd, perfect, ideaal, het (allerl)?beste?", alle, niet (één)leen, geen enkele, niemand, geen een, niets, niks, nergens, nooit, waardeloos, het (allerl)?ergste?, noch, (of)ofwel)(.+ l)of, zwart[-/]wit, zwart of wit, oot
Disqualifying the positive	top,? maar, geweldig,? maar, goed,? maar, ok(elé)?,? maar, niet echt (geweldig)top, niet zo goed, het was niet, niet zo, prima,? maar, acceptabel,? maar, afdoende,? maar
Emotional reasoning	maar ik voel, maar ik heb het gevoel, want ik voel, want ik heb het gevoel, omdat ik voel, omdat ik heb het gevoel, maar het(+ l)voelt, want het(+ l)voelt, omdat het(+ l)voelt, voelt nog steeds
Fortune-telling	ik zal(+ l) niet, we zullen(+ l) niet, j(ijle) (zallzult)(.+ l) niet, z(ijle) zullen(+ l) niet, het zal(+ l)niet, dat zal(+ l)niet, hij zal(+ l)niet, z(elij) zal(+ l)niet
Labeling and mislabeling	ik ben een, hij is een, z(elij) is een, z(elij) zijn een, (het)l't is een, d(ali)t is een, slecht in, waardeloos voor, ik(+ l)nooit, hij(+ l)nooit, z(elij)(.+ l)nooit, w(elij)(.+ l)nooit, j(elij)(.+ l)nooit, (een)tot last, een comple(et)te, een tota(alle), een enorme?, een gro(telot), een loser, een mislukking, een zwak(ke)?, een absolu(t)te, een volslagen, een slechte, een gebroken, een kapotte, een beschadigde?, een hulpelo(oslze), een hopelo(oslze), een incompetent?, een onbekwa(am)me, een giftige?, een toxische?, een lelijke?, een ongewenste?, een ongeerlijke?, een niet geliefd, een ongeliefde?, een waardelo(oslze), een vreselijke?, een afschuwelijke?, een verschrikkelijke?
Magnification and minimization	slechtste?, beste?, niet belangrijk, niet relevant, irrelevant, telt niet, onbelangrijk, maakt niet uit, het enige
Mental filtering	ik zie alleen, ik zie slechts, het enige (w)ld)at ik zie, kan alleen(maar)? denken, (niks)niets goeds?, (niks)niets juist, (compleet)heel) slecht, (compleet)heel) (verkeerd)fout, alleen het (slecht)verkeerde, alleen het slechtste, alleen het ergste, (als)kon) ik maar, ik wou dat ik, als ik nou, (als)kon) het maar, als het nou
Mindreading	iedereen gelooft, iedereen weet, iedereen denkt, iedereen zal(+ l)geloven, iedereen zal(+ l)weten, iedereen zal(+ l)denken, niemand gelooft, niemand weet, niemand denkt, niemand zal(+ l)geloven, niemand zal(+ l)weten, niemand zal(+ l)denken, hij gelooft, hij weet, hij denkt, hij gelooft(+ l)niet, hij weet(+ l)niet, hij denkt(+ l)niet, hij zal(+ l)geloven, hij zal(+ l)weten, hij zal(+ l)denken, hij zal(+ l)niet geloven, hij zal(+ l)niet weten, hij zal(+ l)niet denken, z(ijle) gelooft, z(ijle) weet, z(ijle) denkt, z(ijle) zal(+ l)geloven, z(ijle) zal(+ l)weten, z(ijle) zal(+ l)denken, z(ijle) zal(+ l)niet geloven, z(ijle) zal(+ l)niet weten, z(ijle) zal(+ l)niet denken, z(ijle) geloven, z(ijle) weten, z(ijle) denken, z(ijle) geloven(+ l)niet, z(ijle) weten(+ l)niet, z(ijle) denken(+ l)niet, z(ijle) zullen(+ l)geloven, z(ijle) zullen(+ l)weten, z(ijle) zullen(+ l)denken, z(ijle) zullen(+ l)niet geloven, z(ijle) zullen(+ l)niet weten, z(ijle) zullen(+ l)niet denken, w(ijle) geloven, w(ijle) weten, w(ijle) denken, w(ijle) geloven(+ l)niet, w(ijle) weten(+ l)niet, w(ijle) denken(+ l)niet, w(ijle) zullen(+ l)geloven, w(ijle) zullen(+ l)weten, w(ijle) zullen(+ l)denken, w(ijle) zullen(+ l)niet geloven, w(ijle) zullen(+ l)niet weten, w(ijle) zullen(+ l)niet denken, j(ijle) gelooft, j(ijle) weet, j(ijle) denkt, j(ijle) gelooft(+ l)niet, j(ijle) weet(+ l)niet, j(ijle) denkt(+ l)niet, j(ijle) (zallzult)(.+ l)geloven, j(ijle) (zallzult)(.+ l)weten, j(ijle) (zallzult)(.+ l)denken, j(ijle) (zallzult)(.+ l)niet geloven, j(ijle) (zallzult)(.+ l)niet weten, j(ijle) (zallzult)(.+ l)niet denken
Overgeneralizing	de hele tijd, gebeurt altijd, (steeds)altijd) zo, gebeurt (elk)iedere) keer, compleet", "helemaal, niemand(+ l)oot, z(ijle)(.+ l)allemaal, jullie(+ l)allemaal, ik(+ l)altijd, j(ijle)(.+ l)altijd, hij(+ l)altijd, z(ijle)(.+ l)altijd, ik ben altijd, j(ijle) bent altijd, hij is altijd, z(ijle) is altijd, z(ijle) zijn altijd
Personalizing	(helal)lemaal mijn, omdat ik, omdat mijn, (door)vanwege) mijn, (door)vanwege) mij, ik ben verantwoordelijk, geef mij de schuld, mij de schuld geven, ik veroorzaakte, ik heb(+ l)veroorzaakt, ik voel m(ijle) verantwoordelijk, (helal)lemaal door mij, allemaal mijn schuld, mijn (fout)schuld, mijn verantwoordelijkheid
Should statements	zou, zouden, moet, moeten

Measuring CDS in Chat Messages

Let S denote the set of all chat sessions. Each chat session $s \in S$ in our sample set consists of a collection of k time-ordered messages, $s = \{m_1, m_2, \dots, m_k\}$. Each message m_i in the session is marked by an indicator of the role of who posted that message denoted by r_{m_i} , that is, counselor, help seeker, or system, denoted by o , h , or b , respectively, and the time at which the message was posted (in s from the start of the session), denoted by t_{m_i} .

Additionally, as mentioned, we have defined a set of 265 Dutch CDS n-grams, denoted as C . Each CDS, denoted as c , is intended to represent the lexical building blocks of expressing cognitive distortions. The notation for the collection of CDS within a distortion type is indicated in the column "Notation" of Table 1 (eg, C_c denotes all CDS belonging to the category *Catastrophizing*). We define a CDS matching function $g_C(m_i) \rightarrow \{0,1\}$, which assigns each message m_i to 0 or 1 based on whether that message contains at least 1 CDS that is contained in the set $C \subseteq C$.

Calculating CDS Prevalence

As we are interested in the prevalence of cognitive distortions in messages, we measured the fraction of messages posted that contain a CDS. Thus, we calculated the CDS prevalence for a given set of CDS C , a role r (either help seeker or counselor), and a set of chat session S as the ratio of messages that contain a CDS marker over the total number of messages.

For the NL cohort, S can be seen as the set of timelines of the individuals in the cohort, in which $s \in S$ represents the timeline of an individual that consists of a collection of tweets rather than messages. Finally, for the NL web dataset, S consists of a single element, that is, a collection of sentences rather than messages.

Calculating CDS Prevalence Ratio

We calculated the prevalence ratio between roles as follows:

$$R_C(r_1, r_2, S) = \frac{P_C(r_1, S)}{P_C(r_2, S)}$$

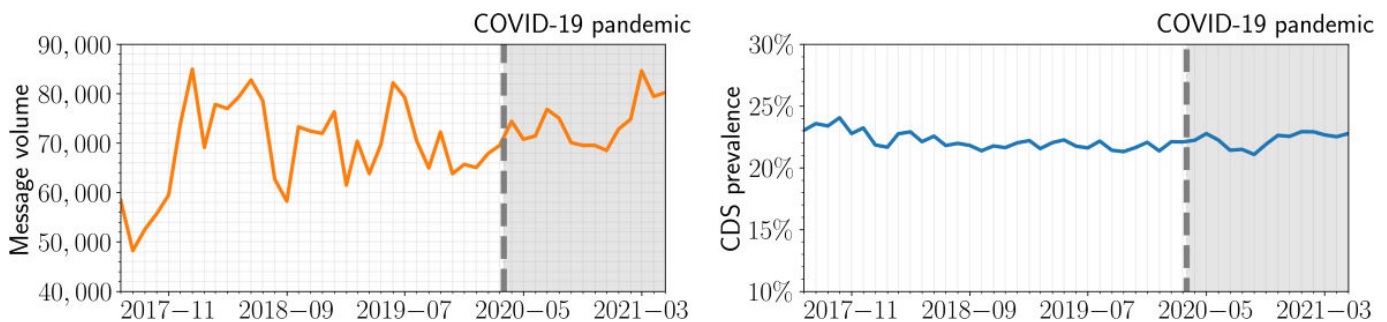
Results

Our analysis of CDS prevalence in the language of 113 help seekers was not diachronic, that is, we compared CDS prevalence for all messages written by a specific cohort

regardless of when the messages were recorded. Therefore, we assessed whether message volume and CDS prevalence remained stable over the time period of our analysis to avoid any bias resulting from our choice of time period. The messages exchanged in the 113 service were recorded in a time interval that spans almost 4 years, roughly from 2017 to 2021, during which changes in both message volume (possibly reflecting lesser or greater prevalence of suicidality) and the nature of the calls (seasonal or topical variation) may have occurred. In particular, the last 2 years of the analysis period include the COVID-19 pandemic (declared a pandemic in March 2020).

As shown in [Figure 1](#), we observed only slight changes in the volume of monthly messages from July 2017 to June 2021; the median monthly message volume is 70,768 (IQR 11,207), and we saw no indications of discontinuities, interruptions, or other changes associated with the start of the COVID-19 pandemic, nor any significant seasonal trends. Importantly, the monthly CDS prevalence (the ratio of monthly CDS matches in messages over the monthly number of messages) showed even less variation, fluctuating in a narrow range around 22.2% (95% CI 21.3-23.5), indicating no significant changes in CDS prevalence over time; hence, there is no indication of topical changes related to the COVID-19 pandemic. The monthly CDS prevalence separated by the CDS type was equally stable over time (see [Figure S1 in Multimedia Appendix 1](#)).

Figure 1. Message volume (number of messages, left panel) and cognitive distortion schemata (CDS) prevalence (%), right panel) for the help seeker cohort over the time period of analysis. The vertical line marks the start of the COVID-19 pandemic in March 2020. No significant changes in message volume nor CDS prevalence (%) were observed for the time period under consideration, including before or after the start of the COVID-19 pandemic.



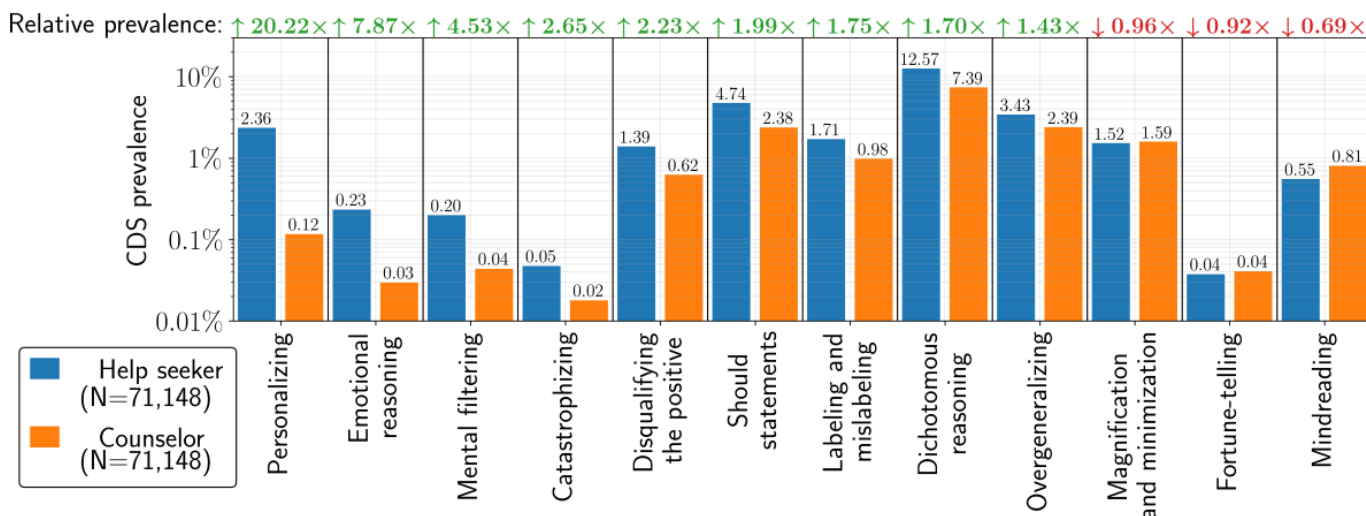
Having verified the absence of significant trends, seasonal changes of message volume nor CDS prevalence over the time period of our analysis, we turned to comparing the cohort-level CDS prevalence between help seekers and counselors across the entire time range. The counselors were used as a control since they are expected to be least affected by suicidality and depression, yet are conversing with help seekers on the same platform, about the same general topic, and in the same language (Dutch).

The overall CDS prevalence between messages entered by both help seekers and counselors, respectively (across the entire time period), showed a pronounced difference of 22.19% and 13.78%, which is mostly concentrated in some CDS types (see [Table 1](#) for denotations), that is, *dichotomous reasoning* ($P_{dr}(h)=0.13$ and $P_{dr}(o)=0.07$), *should statements* ($P_{ss}(h)=0.05$ and $P_{ss}(o)=0.02$), and *overgeneralizing*

($P_o(h)=0.03$ and $P_o(o)=0.02$), for the help seeker and counselor cohorts separately, indicating CDS type-specific effects.

Based on a comparison of CDS prevalence, overall or separated by the CDS type, help seekers exhibited higher CDS prevalence in their language than 113 counselors. To determine the magnitude of this effect by the CDS type, we calculated the *CDS prevalence ratio* by the CDS type between the help seeker and counselor cohorts (CDS prevalence ratio $R = \frac{P(h)}{P(o)}$) as shown in [Figure 2](#). The prevalence ratio between the h and o cohorts can be separated by the CDS type; for example, for CDS *personalizing* types, we can calculate the CDS prevalence ratio between the help seeker and counselor cohorts R_p .

Figure 2. Cognitive distortion schemata (CDS) prevalence for the help seeker and counselor cohorts stratified per category. CDs types are ranked left to right by the magnitude of CDS prevalence ratio. Each set of 2 bars indicates the CDS prevalence P_C for the type C , separated by the help seeker cohort ($P_C(h)$, blue) and the counselor cohort ($P_C(o)$, orange), respectively. The prevalence ratio ($R_C = \frac{P_C(h)}{P_C(o)}$) for each category C is listed at the top with up-and-down arrows indicating whether the CDS of a particular category is either more or less prevalent in the help seeker cohort than the counselor cohort for that category.



As shown in Figure 2, we found that although CDS prevalence is strongly concentrated in some CDS types, the CDS prevalence in the help seeker cohort was significantly larger than CDS prevalence in the counselor cohort for nearly all 12 CDS types, with the exception of “magnification and minimization,” “fortune telling,” and “mindreading.”

In fact, CDS prevalence was 20 times higher in help seeker messages than those of 113 counselors for the *Personalizing* types of CDS; $R_p=20.22\times$, followed by *emotional reasoning* ($R_{er}=7.87\times$), *mental filtering* ($R_p=4.53\times$), *catastrophizing* ($R_c=2.65\times$), and *disqualifying the positive* ($R_{dtp}=2.23\times$).

In addition to the analysis of the differences between the 2 cohorts, we extended our analysis by investigating the difference in CDS prevalence at the level of individual chat sessions. For this analysis, we determined the prevalence of CDS phrases across the messages in a singular chat session for each role (ie, help seeker and counselor) and subsequently analyzed the distribution of this quantity across all chat sessions in the dataset. This analysis sheds more light on interpersonal differences between the individuals across both considered cohorts. Based on this analysis at the level of chat sessions, we found that the mean values of *within-session* CDS prevalence for help seeker and counselor cohorts are 0.228 (SD 0.119) and 0.135 (SD 0.075), respectively. A Bartlett test ($T=14,581, P<.001$) indicated that the CDS prevalence SDs are different for the help seeker and the counselor cohorts. We therefore used the Welch *t* test to find that CDS prevalence was different across the help seeker and counselor cohorts ($t_{120,048}=177, P<.001, d=0.94$).

The language of 113 counselors may vary significantly from that of the general population in an online chat environment and could be expected to have exceptionally low levels of CDS prevalence relative to the general population,

given their training and professional role. Therefore, we also included a comparison with the observed *within-individual* CDS prevalence of a cohort of Twitter users in the NL cohort and a collection of Dutch sentences extracted from the web.

The first was a comparison to the language of a large sample of Twitter (now X) users, since Twitter as a micro-blogging platform was specifically designed for the communication of short, chat-like messages, not dissimilar in format and style from an online chat environment. Using a similar approach to the study by Bathina et al [50], we established a cohort of Twitter users who indicated on their user profile that they lived in Dutch municipalities with more than 100,000 inhabitants in 2020 [55], which we refer to as the NL cohort. For each of the 4829 individuals in this NL cohort, we retrieved a record of up to 3200 of their most recent Tweets, which were subjected to the same CDS prevalence analysis as the messages and sessions of 113 help seekers and counselors.

The second was a comparison with a large corpus of general Dutch text observed online using the January to December 2018 Commoncrawl snapshots from the CC-Net repository [53], referred to as the NL web dataset. This snapshot contains a large collection of Dutch sentences (N=229,914,131), called the CC100-Dutch dataset.

The mean *within-individual* CDS prevalence for the Twitter NL cohort is 0.192 (SD 0.107), meaning that 19.2% of the tweets of the individuals in the NL cohort contain a CDS. Moreover, 20.8% of the Dutch sentences in the NL web dataset contain a CDS. The overall distribution of CDS prevalence for both the chat sessions and *Twitter* timelines is shown in Figure 3. We compared the CDS prevalence across these cohorts with Bartlett and Welch *t* tests, adjusting for multiple comparisons with a Bonferroni correction, the results of which are shown in Table 3.

Figure 3. Within-session cognitive distortion schemata (CDS) prevalence distributions for the help seeker and counselor cohorts (N=71,148 chat sessions), the NL cohort (N=4829 Twitter timelines), and the mean CDS prevalence in the NL web dataset (N=229,914,131 sentences). The panel on the left displays the kernel density estimate based on the observations in each cohort. The 3 panels on the right display the distributions of the observations in each cohort. The dashed vertical line indicates the mean CDS prevalence.

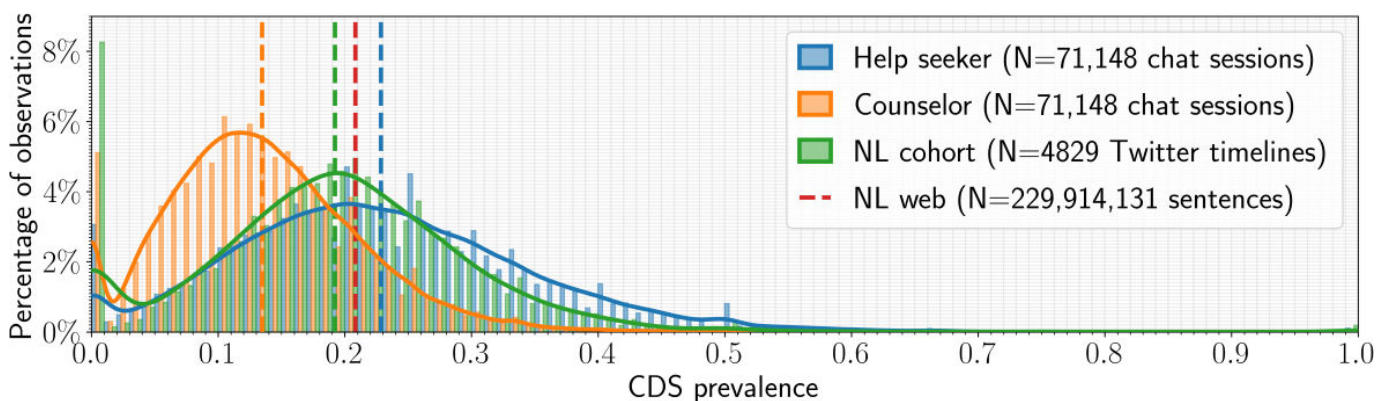


Table 3. Comparing the variance (Bartlett *t* test) and mean (Welch *t* test) of cognitive distortion schemata (CDS) prevalence distributions for (1) help seekers, (2) 113 counselors, and (3) individuals living in the Netherlands on Twitter^a.

Cohort comparison	Bartlett test		Welch <i>t</i> test		
	<i>T</i>	<i>P</i> value	<i>t</i> (<i>df</i>)	<i>P</i> value	Cohen <i>d</i>
Help seeker versus counselor	14,581	<.001	177 (120,048)	<.001	0.94
Help seeker versus NL cohort	93	<.001	22.6 (5667)	<.001	0.031
Counselor versus NL cohort	1415	<.001	-36.7 (5155)	<.001	0.74

^aTo account for multiple hypothesis testing, we applied a Bonferroni correction, that is, $\alpha=0.05/3$ to determine significance levels.

For all comparisons, we could reject the null-hypothesis at the α level given by the Bonferroni correction. The relative differences indicated that the counselor cohort exhibits the lowest CDS prevalence, possibly because of their training and their role as providers of assistance in exchanges with individuals seeking help with suicidality. Although the mean values of CDS prevalence for the NL web, the NL cohort, and the help seeker cohort were very close to each other when examining the distributions, the help seeker cohort had a greater likelihood of individuals with a high CDS prevalence than the NL cohort.

Discussion

Principal Results

We employed a theory-driven approach to investigate the prevalence of expressions of cognitive distortions, that is, maladaptive thought patterns that characterize internalizing disorders, such as depression and anxiety [41-43], in a large collection of chat sessions of a suicide prevention helpline in the NL between help seekers and counselors and 2 baselines extracted from large-scale corpora of Dutch social media and web content.

Our analysis investigated the CDS prevalence in the recorded chat sessions of the Dutch suicide prevention helpline. Our aim was to quantify the levels of cognitive distortions in the language of online help seekers versus 2 controls that include the 113 respondents. Given the comorbidity between suicide and depression [36-38], and the fact that the latter is associated with cognitive distortions [41-

43], we expected high levels of cognitive distortions in the language of those who contact suicide helplines.

The monthly CDS prevalence was quite stable during the entire observation period, even during the 2 years of a devastating worldwide pandemic, which according to some analyses led to topical changes in Dutch suicide helpline conversations (eg, an 808% increase of the topic of “COVID-19”) [61]. This shows that even though the topics that are discussed in the chats change over time, the degree of distorted expressions does not seem to change. This is surprising given the backdrop of a global pandemic. Future work may explore the stability of CDS prevalence across time in online environments further. There are some indications of a minor trend toward increased session volume starting in early 2021. However, this does not seem to affect the monthly CDS prevalence significantly, which is in line with other research studies that showed that the prevalence of high anxiety and depression symptom levels did not increase [62]. This is the case across all separate CDS types (see Figure S1 in Multimedia Appendix 1).

In general, certain types of CDS n-grams are more prevalent than others; the CDS prevalence for *fortune-telling* remained low at 0.04%, while the CDS prevalence for *dichotomous reasoning* was 2 orders of magnitude more prevalent at 12.57% across all chat sessions, as indicated in Figure 2. In general, categories that relate to black-and-white thinking (eg, *dichotomous reasoning* and *overgeneralizing*) were most frequent among both help seekers and counselors, whereas categories that relate to emotion and affect (eg, *catastrophizing*, *emotional reasoning*, and *disqualifying the positive*) were less frequent. This could reflect how some

distortion types are generally more or less common in Dutch or how CDS n-grams of a specific type are worded in the lexicon. In fact, many n-grams contain some of the most frequent Dutch words (eg, “ik” (I) or “ben” (am)), since the CDS lexicon was designed to capture common, context-free patterns of language. It is expected that shorter n-grams occur more frequently in the language than longer n-terms.

We compared the CDS prevalence ratio between help seekers and counselors, since the latter is least likely to be affected by suicidality and depression. We observe that nearly all types of cognitive distortions are significantly more prevalent in help seeker messages than those of the 113 counselors. CDS of the *personalizing*, *emotional reasoning*, and *mental filtering* types are, respectively, 20.22, 7.87, and 4.53 times more prevalent in the help seeker cohort than the counselor cohorts. In fact, 5 categories are used at least twice as frequently in the help seeker cohort as in the counselor cohort, that is, *personalizing*, *emotional reasoning*, *mental filtering*, *catastrophizing*, and *disqualifying the positive*. The fact that the differences for other categories are less pronounced could be explained by counselors mimicking the phrasings of help seekers in their chat in their attempt to build connection with the help seeker, a phenomenon called lexical entrainment [63]. The CDS n-grams have in previous research been shown not to have significant valence loadings [50] nor is their prevalence in the language of individuals with depression driven by the presence of personal pronouns. Hence, the expected higher level of negativity or self-referential nature of the help seekers' messages is not likely to explain the higher prevalence of the set of CDS n-grams in the language of help seekers.

Our analysis of the *within*-session CDS prevalence shows clear differences in CDS prevalence between the help seeker and counselor cohorts, with the help seeker cohort having significantly a higher mean CDS prevalence than the counselor cohort. This may be expected because the counselor cohort has been professionally trained to communicate with help seekers, which possibly included the avoidance of cognitive distortions in their communications with help seekers. To verify the degree to which counselors exhibit exceptionally low levels of cognitive distortions in their language, we compared the prevalence of CDS in their language and that of the help seekers to a cohort of Dutch individuals (referred to as the NL cohort) on Twitter, since the latter as the microblogging platform may closely approximate the format and vernacular of online chat helplines.

As expected, the counselor cohort has significantly a lower mean CDS prevalence than the NL cohort, that is, their language contains fewer markers of cognitive distortions than a random sample of Dutch Twitter users. As mentioned, this lower usage of CDS in the counselor cohort is likely a result of the training they have received in handling these chat sessions at 113. Another fact that stands out in Figure 3 is the difference in variability between the help seeker and the counselor cohorts, which indicates that there are significant differences between individual chat sessions and that the 113 service is reaching a wide audience of individuals

with varying degrees of urgency and the need of assistance. These results highlight the usability of the CDS annotations in the chat sessions to further analyze these conversations, for example, by providing the counselors with real-time feedback during chat sessions, or by highlighting important changes in language during the sessions that could give an indication of the perceived quality of the conversation.

Our data were collected by an anonymous helpline, which in the context of providing urgent, low-threshold care did not collect information on neurodevelopmental diagnoses. Certain neurodevelopmental conditions have been associated with differences in cognitive styles, which may relate to variations in the expression of cognitive distortions. Additionally, some neurodivergent populations have been shown to experience higher levels of suicidal ideation and suicide attempts [64, 65]. However, the relationship between such differences and the prevalence of cognitive distortions remains insufficiently characterized. Future research should examine how these individual factors may interact with the prevalence of cognitive distortions in the context of suicide helplines.

Nevertheless, we have converging evidence across 4 different online samples (113 help seekers, 113 counselors, a large cohort of random Dutch Twitter users, and a large collection of Dutch online texts), indicating that the prevalence of cognitive distortion markers in the language of 113 help seekers is exceptionally high.

Limitations

Several limitations of our approach should be considered.

Constructing a Good Baseline for the 113 Chat Platform Is Challenging

It is difficult to gauge the magnitude of CDS prevalence in the language of the help seekers in practice without a proper baseline or counterfactual. Our sample does not include individuals contacting the same helpline without suicidal ideation or intent while the 113 counselors are trained professionals who may avoid expressions of cognitive distortions in their language or responses to help seekers. Although the format and vernacular of messages on Twitter as a microblogging platform may approximate that of chat environments, it may still fall short in when compared to the 113 chat platform. For instance, the focus of a microblogging platform is not on having conversations as opposed to a helpline, nor will it discuss the same topics.

CDS Annotations Do Not Necessarily Equate to Cognitive Distortion

We emphasize that not all uses of CDS n-grams reflect actual cognitive distortions. Although the set of CDS n-grams was specifically designed to capture the expression of common types of cognitive distortions by experts in CBT and shown to capture differences in the language between individuals with depression and a random sample, these phrases can be part of nondepressed expressions, possibly leading to false positives. However, such false positives would be equally likely in either one of the examined cohorts. Furthermore,

the CDS lexicon does not exhaustively cover all possible expressions of cognitive distortions in the language as it contains 241 n-grams. This may cause false negatives where a cognitive distortion may have been expressed in a statement but was left unrecognized by the CDS lexicon. On balance, the presence of a CDS does not provide proof positive evidence of the presence of maladaptive thought patterns, but consistently observing CDS prevalence across cohorts allows for a fair comparison between cohorts. Future work should explore the validity of CDS occurrence as an indicator of cognitive distortions by comparing them to human annotations.

Cognitive Distortions Do Not Necessarily Map One-to-One to Specific Categories

The CDS n-grams are designed to represent 1 type of cognitive distortion among 12 commonly discussed types in the literature. However, in practice, CDS n-grams of different types can occur in the same message and reflect overlapping types of cognitive distortions. Furthermore, the specific types discussed in the literature are generally used to illustrate the concept of cognitive distortions, without necessarily making claims about the component structure of the concept of cognitive distortions, which may rather all belong to the same class of rigid, unrealistic, and exaggerated thinking. As a result, the observed CDS prevalence by category likely underestimates true prevalence. Based on the distribution of *within-session* CDS prevalence of both cohorts, this underestimation reduces the relative difference in the CDS prevalence of a given category, reducing the observed effects instead of exaggerating them.

Ethical Reflections

Given the sensitive nature of both the data and the possible applications of our considered technique, there are several ethical reflections that we want to address explicitly. First,

given the sensitive nature of the content of the chat sessions, all data were processed in the aggregate, without personal identifiers, ensuring the anonymity of all analyses and research outcomes. Next, we stress that our aim is not to diagnose individuals with cognitive distortions or suicidality nor is it our intention to underpin replacements of human counselors by automated chatbots [66,67]. Based on our current results, we believe that our research has the potential to provide additional guidance or support to the counselors of the suicide prevention helpline by highlighting important changes in language that are known to be implicated in the development and recovery trajectories of internalizing disorders. However, careful considerations must be made before such techniques are put into practice [67,68]. Future work will focus on more accurate methods to detect cognitive distortions in language, such as the use of more sophisticated natural language processing and artificial intelligence, including ground truth obtained from human raters. Our analysis did not investigate changes in language over the course of a chat session as the conversation between the help seeker and the counselor unfolds, which may provide indicators of the effectiveness of the exchange. Future research should examine whether the changes of CDS prevalence during a conversation, in subsequent utterances, may shape its perceived quality and effectiveness.

Conclusions

We presented converging evidence across 4 different online samples (113 help seekers, 113 counselors, a large cohort of random Dutch Twitter users, and a large collection of Dutch online texts), indicating that the prevalence of cognitive distortion markers in the language of 113 help seekers is exceptionally high. Our results point to the value of lexical analysis as an instrument to detect the cognitive and lexical markers of suicidality.

Acknowledgments

The authors thank Wilco Janssen, Jeroen Gomes, Claudi Bockting, and Marcus Huibers for their contributions in translating the cognitive distortion schemata (CDS) from English to Dutch. The authors stated that no generative artificial intelligence tools were used.

Funding

This research was supported, in part, by the National Institute of Health (NIH, grant numbers R01-CA279679-01A1 and 1R21HD108052-01A1) and the National Institute of Mental Health (NIMH, grant number 1R01MH135502-01A1). The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies above.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Authors' Contributions

MtT and JB conceptualized the analysis and designed the methodology. MtT constructed the dataset and performed data analysis. MtT, SM, RG, and JB wrote and reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The variations in frequency and prevalence of CDS markers in both the help seeker and operator cohorts. Moreover, it presents an analysis that compares different types of help seekers (individuals with suicide ideation vs friends/next of kin) that reach out to the 113 helpline as a control group.

[[DOCX File \(Microsoft Word File\), 2099 KB-Multimedia Appendix 1](#)]

References

1. Suicide worldwide in 2019: global health estimates. World Health Organization; 2021. URL: <https://iris.who.int/server/api/core/bitstreams/3bd4ac79-4347-420e-b675-948d36ab3d90/content> [Accessed 2026-03-31]
2. Fanslow J, Coggan C, Miller B, Norton R. The economic cost of homicide in New Zealand. *Soc Sci Med*. Oct 1997;45(7):973-977. [doi: [10.1016/s0277-9536\(97\)00006-3](https://doi.org/10.1016/s0277-9536(97)00006-3)] [Medline: [9257390](https://pubmed.ncbi.nlm.nih.gov/9257390/)]
3. Corso PS, Mercy JA, Simon TR, Finkelstein EA, Miller TR. Medical costs and productivity losses due to interpersonal and self-directed violence in the United States. *Am J Prev Med*. Jun 2007;32(6):474-482. [doi: [10.1016/j.amepre.2007.02.010](https://doi.org/10.1016/j.amepre.2007.02.010)] [Medline: [17533062](https://pubmed.ncbi.nlm.nih.gov/17533062/)]
4. Kennelly B. The economic cost of suicide in Ireland. *Crisis*. 2007;28(2):89-94. [doi: [10.1027/0227-5910.28.2.89](https://doi.org/10.1027/0227-5910.28.2.89)] [Medline: [17722690](https://pubmed.ncbi.nlm.nih.gov/17722690/)]
5. Kinchin I, Doran CM. The cost of youth suicide in Australia. *Int J Environ Res Public Health*. Apr 4, 2018;15(4):672. [doi: [10.3390/ijerph15040672](https://doi.org/10.3390/ijerph15040672)] [Medline: [29617305](https://pubmed.ncbi.nlm.nih.gov/29617305/)]
6. Shepard DS, Gurewich D, Lwin AK, Reed GA Jr, Silverman MM. Suicide and suicidal attempts in the United States: costs and policy implications. *Suicide & Life Threat Behav*. Jun 2016;46(3):352-362. [doi: [10.1111/sltb.12225](https://doi.org/10.1111/sltb.12225)]
7. Pil L, Pauwels K, Muijzers E, Portzky G, Annemans L. Cost-effectiveness of a helpline for suicide prevention. *J Telemed Telecare*. Jul 2013;19(5):273-281. [doi: [10.1177/1357633X13495487](https://doi.org/10.1177/1357633X13495487)] [Medline: [24163237](https://pubmed.ncbi.nlm.nih.gov/24163237/)]
8. Looijmans M, Elzinga E, Popma A, van Bergen D, Gilissen R, Mérelle S. Understanding the needs and perspectives of young adults with recent suicidal ideation: insights for suicide prevention. *Front Child Adolesc Psychiatry*. 2024;3:1376872. [doi: [10.3389/frcha.2024.1376872](https://doi.org/10.3389/frcha.2024.1376872)] [Medline: [39816575](https://pubmed.ncbi.nlm.nih.gov/39816575/)]
9. Key substance use and mental health indicators in the United States: results from the 2018 National Survey on Drug Use and Health. Substance Abuse and Mental Health Services Administration; 2019. URL: <https://mambh-web.files.svcdcdn.com/production/files/NSDUHNationalFindingsReport2018.pdf?dm=1615222948> [Accessed 2026-03-31]
10. Ivey-Stephenson AZ, Demissie Z, Crosby AE, et al. Suicidal ideation and behaviors among high school students - Youth Risk Behavior Survey, United States, 2019. *MMWR Suppl*. Aug 21, 2020;69(1):47-55. [doi: [10.15585/mmwr.su6901a6](https://doi.org/10.15585/mmwr.su6901a6)] [Medline: [32817610](https://pubmed.ncbi.nlm.nih.gov/32817610/)]
11. Patel V, Saxena S, Lund C, et al. The Lancet commission on global mental health and sustainable development. *The Lancet*. Oct 2018;392(10157):1553-1598. [doi: [10.1016/S0140-6736\(18\)31612-X](https://doi.org/10.1016/S0140-6736(18)31612-X)]
12. Mokkenstorm JK, Eikelenboom M, Huisman A, et al. Evaluation of the 113 online suicide prevention crisis chat service: outcomes, helper behaviors and comparison to telephone hotlines. *Suicide Life Threat Behav*. Jun 2017;47(3):282-296. [doi: [10.1111/sltb.12286](https://doi.org/10.1111/sltb.12286)] [Medline: [27539122](https://pubmed.ncbi.nlm.nih.gov/27539122/)]
13. Hoffberg AS, Stearns-Yoder KA, Brenner LA. The effectiveness of crisis line services: a systematic review. *Front Public Health*. 2019;7:399. [doi: [10.3389/fpubh.2019.00399](https://doi.org/10.3389/fpubh.2019.00399)] [Medline: [32010655](https://pubmed.ncbi.nlm.nih.gov/32010655/)]
14. Zalsman G, Hawton K, Wasserman D, et al. Suicide prevention strategies revisited: 10-year systematic review. *Lancet Psychiatry*. Jul 2016;3(7):646-659. [doi: [10.1016/S2215-0366\(16\)30030-X](https://doi.org/10.1016/S2215-0366(16)30030-X)] [Medline: [27289303](https://pubmed.ncbi.nlm.nih.gov/27289303/)]
15. Sindahl TN, Côte LP, Dargis L, Mishara BL, Bechmann Jensen T. Texting for help: processes and impact of text counseling with children and youth with suicide ideation. *Suicide Life Threat Behav*. Oct 2019;49(5):1412-1430. [doi: [10.1111/sltb.12531](https://doi.org/10.1111/sltb.12531)] [Medline: [30468267](https://pubmed.ncbi.nlm.nih.gov/30468267/)]
16. Mazzer K, O'Riordan M, Woodward A, Rickwood D. A systematic review of user expectations and outcomes of crisis support services. *Crisis*. Nov 2021;42(6):465-473. [doi: [10.1027/0227-5910/a000745](https://doi.org/10.1027/0227-5910/a000745)] [Medline: [33275048](https://pubmed.ncbi.nlm.nih.gov/33275048/)]
17. Mathieu SL, Uddin R, Brady M, et al. Systematic review: the state of research into youth helplines. *J Am Acad Child Adolesc Psychiatry*. Oct 2021;60(10):1190-1233. [doi: [10.1016/j.jaac.2020.12.028](https://doi.org/10.1016/j.jaac.2020.12.028)]
18. Brody C, Star A, Tran J. Chat-based hotlines for health promotion: a systematic review. *Mhealth*. 2020;6:36. [doi: [10.21037/mhealth-2019-di-13](https://doi.org/10.21037/mhealth-2019-di-13)] [Medline: [33437833](https://pubmed.ncbi.nlm.nih.gov/33437833/)]
19. Gould MS, Chowdhury S, Lake AM, et al. National Suicide Prevention Lifeline crisis chat interventions: evaluation of chatters' perceptions of effectiveness. *Suicide Life Threat Behav*. Dec 2021;51(6):1126-1137. [doi: [10.1111/sltb.12795](https://doi.org/10.1111/sltb.12795)] [Medline: [34331471](https://pubmed.ncbi.nlm.nih.gov/34331471/)]
20. Pauwels K, De Jaegere E, Vanderreydt P, Aerts S, Vande Gaer E, Portzky G. Assessing a suicide prevention helpline's impact on caller crisis level and suicidality. *Arch Suicide Res*. 2025;29(3):621-636. [doi: [10.1080/13811118.2024.2394666](https://doi.org/10.1080/13811118.2024.2394666)] [Medline: [39177448](https://pubmed.ncbi.nlm.nih.gov/39177448/)]

21. O'Dea B, Larsen ME, Batterham PJ, Calear AL, Christensen H. A linguistic analysis of suicide-related Twitter posts. *Crisis*. Sep 2017;38(5):319-329. [doi: [10.1027/0227-5910/a000443](https://doi.org/10.1027/0227-5910/a000443)] [Medline: [28228065](https://pubmed.ncbi.nlm.nih.gov/28228065/)]
22. Lao C, Lane J, Suominen H. Analyzing suicide risk from linguistic features in social media: evaluation study. *JMIR Form Res*. Aug 30, 2022;6(8):e35563. [doi: [10.2196/35563](https://doi.org/10.2196/35563)] [Medline: [36040781](https://pubmed.ncbi.nlm.nih.gov/36040781/)]
23. Yeskuatov E, Chua SL, Foo LK. Detecting suicidal ideations in online forums with textual and psycholinguistic features. *Appl Sci (Basel)*. 2024;14(21):9911. [doi: [10.3390/app14219911](https://doi.org/10.3390/app14219911)]
24. Ji S, Yu CP, Fung S fu, Pan S, Long G. Supervised learning for suicidal ideation detection in online user content. *Complexity*. Jan 2018;2018(1):6157249. [doi: [10.1155/2018/6157249](https://doi.org/10.1155/2018/6157249)]
25. Wiltsey Stirman S, Pennebaker JW. Word use in the poetry of suicidal and nonsuicidal poets. *Psychosom Med*. Jul 2001;63(4):517-522. [doi: [10.1097/00006842-200107000-00001](https://doi.org/10.1097/00006842-200107000-00001)]
26. De Choudhury M, Kiciman E, Dredze M, Coppersmith G, Kumar M. Discovering shifts to suicidal ideation from mental health content in social media. Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, California, USA. URL: <https://dl.acm.org/doi/abs/10.1145/2858036.2858207> [Accessed 2026-03-31] [doi: [10.1145/2858036.2858207](https://doi.org/10.1145/2858036.2858207)]
27. De Choudhury M, Kiciman E. The language of social support in social media and its effect on suicidal ideation risk. *Proc Int AAAI Conf Weblogs Soc Media*. May 2017;11(1):32-41. [doi: [10.1609/icwsm.v11i1.14891](https://doi.org/10.1609/icwsm.v11i1.14891)]
28. Feldhege J, Wolf M, Moessner M, Bauer S. Psycholinguistic changes in the communication of adolescent users in a suicidal ideation online community during the COVID-19 pandemic. *Eur Child Adolesc Psychiatry*. Jun 2023;32(6):975-985. [doi: [10.1007/s00787-022-02067-7](https://doi.org/10.1007/s00787-022-02067-7)] [Medline: [36018514](https://pubmed.ncbi.nlm.nih.gov/36018514/)]
29. Egnoto MJ, Griffin DJ. Analyzing language in suicide notes and legacy tokens. *Crisis*. Mar 2016;37(2):140-147. [doi: [10.1027/0227-5910/a000363](https://doi.org/10.1027/0227-5910/a000363)] [Medline: [27232429](https://pubmed.ncbi.nlm.nih.gov/27232429/)]
30. Teixeira AS, Talaga S, Swanson TJ, Stella M. Revealing semantic and emotional structure of suicide notes with cognitive network science. *Sci Rep*. Sep 30, 2021;11(1):19423. [doi: [10.1038/s41598-021-98147-w](https://doi.org/10.1038/s41598-021-98147-w)] [Medline: [34593826](https://pubmed.ncbi.nlm.nih.gov/34593826/)]
31. Leavitt J, Hong JH, Walker RL. Paradoxical positivity: suicide notes use less distressed language than blogs about depression, suicidal thoughts, and even cooking. *Suicide Life Threat Behav*. Oct 2021;51(5):1005-1014. [doi: [10.1111/sltb.12786](https://doi.org/10.1111/sltb.12786)] [Medline: [34184305](https://pubmed.ncbi.nlm.nih.gov/34184305/)]
32. Pennebaker JW, Boyd RL, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. University of Texas at Austin; 2015. URL: <https://repositories.lib.utexas.edu/server/api/core/bitstreams/b0d26dcf-2391-4701-88d0-3cf50ebee697/content> [Accessed 2026-03-31]
33. Brancu M, Jobes D, Wagner BM, Greene JA, Fratto TA. Are there linguistic markers of suicidal writing that can predict the course of treatment? A repeated measures longitudinal analysis. *Arch Suicide Res*. Jul 2, 2016;20(3):438-450. [doi: [10.1080/13811118.2015.1040935](https://doi.org/10.1080/13811118.2015.1040935)] [Medline: [26219609](https://pubmed.ncbi.nlm.nih.gov/26219609/)]
34. Salmi S, Mérelle S, Gilissen R, Brinkman WP. Content-based recommender support system for counselors in a suicide prevention chat helpline: design and evaluation study. *J Med Internet Res*. Jan 7, 2021;23(1):e21690. [doi: [10.2196/21690](https://doi.org/10.2196/21690)] [Medline: [33410755](https://pubmed.ncbi.nlm.nih.gov/33410755/)]
35. Janssen W, van Raak J, van der Lucht Y, van Ballegooijen W, Mérelle S. Can outcomes of a chat-based suicide prevention helpline be improved by training counselors in motivational interviewing? A non-randomized controlled trial. *Front Digit Health*. 2022;4:871841. [doi: [10.3389/fdgth.2022.871841](https://doi.org/10.3389/fdgth.2022.871841)] [Medline: [35801118](https://pubmed.ncbi.nlm.nih.gov/35801118/)]
36. Wiebenga JXM, Dickhoff J, Mérelle SYM, et al. Prevalence, course, and determinants of suicide ideation and attempts in patients with a depressive and/or anxiety disorder: a review of NESDA findings. *J Affect Disord*. Mar 15, 2021;283:267-277. [doi: [10.1016/j.jad.2021.01.053](https://doi.org/10.1016/j.jad.2021.01.053)] [Medline: [33571797](https://pubmed.ncbi.nlm.nih.gov/33571797/)]
37. Bolton JM, Pagura J, Enns MW, Grant B, Sareen J. A population-based longitudinal study of risk factors for suicide attempts in major depressive disorder. *J Psychiatr Res*. Oct 2010;44(13):817-826. [doi: [10.1016/j.jpsychires.2010.01.003](https://doi.org/10.1016/j.jpsychires.2010.01.003)] [Medline: [20122697](https://pubmed.ncbi.nlm.nih.gov/20122697/)]
38. Baldessarini RJ, Tondo L, Pinna M, Nuñez N, Vázquez GH. Suicidal risk factors in major affective disorders. *Br J Psychiatry*. Jul 11, 2019;215(4):1-6. [doi: [10.1192/bjp.2019.167](https://doi.org/10.1192/bjp.2019.167)] [Medline: [31292010](https://pubmed.ncbi.nlm.nih.gov/31292010/)]
39. Eichstaedt JC, Smith RJ, Merchant RM, et al. Facebook language predicts depression in medical records. *Proc Natl Acad Sci U S A*. Oct 30, 2018;115(44):11203-11208. [doi: [10.1073/pnas.1802331115](https://doi.org/10.1073/pnas.1802331115)] [Medline: [30322910](https://pubmed.ncbi.nlm.nih.gov/30322910/)]
40. Rude S, Gortner EM, Pennebaker J. Language use of depressed and depression-vulnerable college students. *Cogn Emot*. Dec 2004;18(8):1121-1133. [doi: [10.1080/02699930441000030](https://doi.org/10.1080/02699930441000030)]
41. Beck AT. Thinking and depression. i. Idiosyncratic content and cognitive distortions. *Arch Gen Psychiatry*. Oct 1963;9(4):324-333. [doi: [10.1001/archpsyc.1963.01720160014002](https://doi.org/10.1001/archpsyc.1963.01720160014002)] [Medline: [14045261](https://pubmed.ncbi.nlm.nih.gov/14045261/)]
42. Beck AT. Thinking and depression. *Arch Gen Psychiatry*. Jun 1, 1964;10(6):561-571. [doi: [10.1001/archpsyc.1964.01720240015003](https://doi.org/10.1001/archpsyc.1964.01720240015003)]

43. Beck JS. Cognitive Behavior Therapy: Basics and Beyond. 2nd ed. Guilford Press; 2011. URL: <https://www.amazon.in/Cognitive-Behavior-Therapy-Second-Basics/dp/1609185048> [Accessed 2026-03-31] ISBN: 978-1-60918-504-6
44. Hawton K, Witt KG, Taylor Salisbury TL, et al. Interventions for self-harm in children and adolescents. *Cochrane Database Syst Rev*. Dec 21, 2015;2015(12):CD012013. [doi: [10.1002/14651858.CD012013](https://doi.org/10.1002/14651858.CD012013)] [Medline: [26688129](https://pubmed.ncbi.nlm.nih.gov/26688129/)]
45. Zhou X, Teng T, Zhang Y, et al. Comparative efficacy and acceptability of antidepressants, psychotherapies, and their combination for acute treatment of children and adolescents with depressive disorder: a systematic review and network meta-analysis. *Lancet Psychiatry*. Jul 2020;7(7):581-601. [doi: [10.1016/S2215-0366\(20\)30137-1](https://doi.org/10.1016/S2215-0366(20)30137-1)] [Medline: [32563306](https://pubmed.ncbi.nlm.nih.gov/32563306/)]
46. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Can J Psychiatry*. Jul 2013;58(7):376-385. [doi: [10.1177/070674371305800702](https://doi.org/10.1177/070674371305800702)] [Medline: [23870719](https://pubmed.ncbi.nlm.nih.gov/23870719/)]
47. Karyotaki E, Efthimiou O, Miguel C, et al. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. *JAMA Psychiatry*. Apr 1, 2021;78(4):361-371. [doi: [10.1001/jamapsychiatry.2020.4364](https://doi.org/10.1001/jamapsychiatry.2020.4364)] [Medline: [33471111](https://pubmed.ncbi.nlm.nih.gov/33471111/)]
48. Lorenzo-Luaces L. The evidence for cognitive behavioral therapy. *JAMA*. Feb 27, 2018;319(8):831-832. [doi: [10.1001/jama.2017.20826](https://doi.org/10.1001/jama.2017.20826)] [Medline: [29486024](https://pubmed.ncbi.nlm.nih.gov/29486024/)]
49. Wu H, Lu L, Qian Y, et al. The significance of cognitive-behavioral therapy on suicide: an umbrella review. *J Affect Disord*. Nov 15, 2022;317:142-148. [doi: [10.1016/j.jad.2022.08.067](https://doi.org/10.1016/j.jad.2022.08.067)] [Medline: [36041581](https://pubmed.ncbi.nlm.nih.gov/36041581/)]
50. Bathina KC, Ten Thij M, Lorenzo-Luaces L, Rutter LA, Bollen J. Individuals with depression express more distorted thinking on social media. *Nat Hum Behav*. Apr 2021;5(4):458-466. [doi: [10.1038/s41562-021-01050-7](https://doi.org/10.1038/s41562-021-01050-7)] [Medline: [33574604](https://pubmed.ncbi.nlm.nih.gov/33574604/)]
51. Bollen J, Ten Thij M, Breithaupt F, et al. Historical language records reveal a surge of cognitive distortions in recent decades. *Proc Natl Acad Sci U S A*. Jul 27, 2021;118(30):e2102061118. [doi: [10.1073/pnas.2102061118](https://doi.org/10.1073/pnas.2102061118)] [Medline: [34301899](https://pubmed.ncbi.nlm.nih.gov/34301899/)]
52. Salmi S, Mérelle S, Gilissen R, van der Mei R, Bhulai S. The most effective interventions for classification model development to predict chat outcomes based on the conversation content in online suicide prevention chats: machine learning approach. *JMIR Ment Health*. Sep 26, 2024;11:e57362. [doi: [10.2196/57362](https://doi.org/10.2196/57362)] [Medline: [39326039](https://pubmed.ncbi.nlm.nih.gov/39326039/)]
53. Wenzek G, Lachaux MA, Conneau A, Chaudhary V, Guzmán F. CCNet: extracting high quality monolingual datasets from web crawl data. Presented at: Proceedings of the Twelfth Language Resources and Evaluation Conference; May 11-16, 2020; Marseille, France. URL: <https://aclanthology.org/2020.lrec-1.494/> [Accessed 2026-03-31]
54. Chen E, Lerman K, Ferrara E. Tracking social media discourse about the COVID-19 pandemic: development of a public coronavirus Twitter data set. *JMIR Public Health Surveill*. May 29, 2020;6(2):e19273. [doi: [10.2196/19273](https://doi.org/10.2196/19273)] [Medline: [32427106](https://pubmed.ncbi.nlm.nih.gov/32427106/)]
55. Population on January 1st. CBS Statline; 2020. URL: <https://opendata.cbs.nl/#/CBS/nl/dataset/70072ned/table?ts=1775640759560> [Accessed 2026-04-08]
56. Wang Z, Hale S, Adelani DI, et al. Demographic inference and representative population estimates from multilingual social media data. Presented at: Proceedings of the 2019 World Wide Web Conference (WWW '19); May 13-17, 2019; San Francisco, CA, USA. URL: <https://publikationen.sulb.uni-saarland.de/bitstream/20.500.11880/33357/1/p2056-wang%20%281%29.pdf> [Accessed 2026-03-31] [doi: [10.1145/3308558.3313684](https://doi.org/10.1145/3308558.3313684)]
57. Tauscher JS, Lybarger K, Ding X, et al. Automated detection of cognitive distortions in text exchanges between clinicians and people with serious mental illness. *Psychiatr Serv*. Apr 1, 2023;74(4):407-410. [doi: [10.1176/appi.ps.202100692](https://doi.org/10.1176/appi.ps.202100692)] [Medline: [36164769](https://pubmed.ncbi.nlm.nih.gov/36164769/)]
58. Simms T, Ramstedt C, Rich M, Richards M, Martinez T, Giraud-Carrier C. Detecting cognitive distortions through machine learning text analytics. Presented at: 2017 IEEE International Conference on Healthcare Informatics (ICHI); Aug 23-26, 2017; Park City, UT, USA. URL: <https://ieeexplore.ieee.org/abstract/document/8031202> [Accessed 2026-03-31] [doi: [10.1109/ICHI.2017.39](https://doi.org/10.1109/ICHI.2017.39)]
59. Ernala SK, Birnbaum ML, Candan KA, et al. Methodological gaps in predicting mental health states from social media: triangulating diagnostic signals. Presented at: CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; May 4-9, 2019; Glasgow, Scotland, UK. URL: <http://www.munmund.net/pubs/CHI19-MethodGaps.pdf> [Accessed 2026-03-31] [doi: [10.1145/3290605.3300364](https://doi.org/10.1145/3290605.3300364)]
60. Shickel B, Siegel S, Heesacker M, Benton S, Rashidi P. Automatic detection and classification of cognitive distortions in mental health text. Presented at: 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE); Oct 26-28, 2020; Cincinnati, OH, USA. URL: <https://ieeexplore.ieee.org/abstract/document/9288097> [Accessed 2026-03-31] [doi: [10.1109/BIBE50027.2020.00052](https://doi.org/10.1109/BIBE50027.2020.00052)]
61. Salmi S, Mérelle S, Gilissen R, van der Mei R, Bhulai S. Detecting changes in help seeker conversations on a suicide prevention helpline during the COVID-19 pandemic: in-depth analysis using encoder representations from transformers. *BMC Public Health*. Mar 18, 2022;22(1):530. [doi: [10.1186/s12889-022-12926-2](https://doi.org/10.1186/s12889-022-12926-2)] [Medline: [35300638](https://pubmed.ncbi.nlm.nih.gov/35300638/)]

62. van der Velden PG, Contino C, Das M, van Loon P, Bosmans MWG. Anxiety and depression symptoms, and lack of emotional support among the general population before and during the COVID-19 pandemic. a prospective national study on prevalence and risk factors. *J Affect Disord*. Dec 1, 2020;277:540-548. [doi: [10.1016/j.jad.2020.08.026](https://doi.org/10.1016/j.jad.2020.08.026)] [Medline: [32889378](https://pubmed.ncbi.nlm.nih.gov/32889378/)]
63. Brennan SE, Clark HH. Conceptual pacts and lexical choice in conversation. *J Exp Psychol Learn Mem Cogn*. Nov 1996;22(6):1482-1493. [doi: [10.1037//0278-7393.22.6.1482](https://doi.org/10.1037//0278-7393.22.6.1482)] [Medline: [8921603](https://pubmed.ncbi.nlm.nih.gov/8921603/)]
64. van Bentum J, Sijbrandij M, Huibers M, Begeer S. Occurrence and predictors of lifetime suicidality and suicidal ideation in autistic adults. *Autism*. Sep 2024;28(9):2282-2294. [doi: [10.1177/13623613231225901](https://doi.org/10.1177/13623613231225901)] [Medline: [38340029](https://pubmed.ncbi.nlm.nih.gov/38340029/)]
65. Balt E, Mérelle S, Vrinzen S, et al. Mixed-methods psychosocial autopsy study of suicide in young and middle-aged individuals in the Netherlands. *BMJ Public Health*. 2025;3(2):e001651. [doi: [10.1136/bmjph-2024-001651](https://doi.org/10.1136/bmjph-2024-001651)] [Medline: [41306786](https://pubmed.ncbi.nlm.nih.gov/41306786/)]
66. Abd-Alrazaq AA, Rababeh A, Alajlani M, Bewick BM, Househ M. Effectiveness and safety of using chatbots to improve mental health: systematic review and meta-analysis. *J Med Internet Res*. Jul 13, 2020;22(7):e16021. [doi: [10.2196/16021](https://doi.org/10.2196/16021)] [Medline: [32673216](https://pubmed.ncbi.nlm.nih.gov/32673216/)]
67. Sedlakova J, Trachsel M. Conversational artificial intelligence in psychotherapy: a new therapeutic tool or agent? *Am J Bioeth*. May 2023;23(5):4-13. [doi: [10.1080/15265161.2022.2048739](https://doi.org/10.1080/15265161.2022.2048739)] [Medline: [35362368](https://pubmed.ncbi.nlm.nih.gov/35362368/)]
68. Reddy S. Evaluating large language models for use in healthcare: a framework for translational value assessment. *Informatics Med Unlocked*. 2023;41:101304. [doi: [10.1016/j.imu.2023.101304](https://doi.org/10.1016/j.imu.2023.101304)]

Abbreviations

- CBT:** cognitive behavioral therapy
- CDC:** Centers for Disease Control and Prevention
- CDS:** cognitive distortion schemata
- WHO:** World Health Organization

Edited by John Torous; peer-reviewed by Alex Markovits, Edmond Appiah; submitted 24.Jul.2025; final revised version received 19.Feb.2026; accepted 09.Mar.2026; published 28.Apr.2026

Please cite as:

ten Thij M, Mérelle S, Gilissen R, Bollen J

Prevalence of Cognitive Distortion Markers in a Suicide Prevention Chat Service: Mixed Methods Study

JMIR Ment Health 2026;13:e81213

URL: <https://mental.jmir.org/2026/1/e81213>

doi: [10.2196/81213](https://doi.org/10.2196/81213)

© Marijn ten Thij, Saskia Mérelle, Renske Gilissen, Johan Bollen. Originally published in JMIR Mental Health (<https://mental.jmir.org>), 28.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.