Published in final edited form as:

Health Commun. 2018 September ; 33(9): 1184–1193. doi:10.1080/10410236.2017.1350906.

What Do You Say Before You Relapse? How Language Use in a Peer-to-Peer Online Discussion Forum Predicts Risky Drinking Among Those in Recovery

Rachel Kornfield,
School of Journalism and Mass Communication, University of Wisconsin-Madison

Catalina L. Toma,
Department of Communication Arts, University of Wisconsin-Madison

Dhavan V. Shah,
School of Journalism and Mass Communication, University of Wisconsin-Madison

Tae Joon Moon, and
School of Journalism and Mass Communication, University of Wisconsin-Madison

David H. Gustafson
Department of Industrial and Systems Engineering, University of Wisconsin-Madison

Abstract

Increasingly, individuals with alcohol use disorder (AUD) seek and provide support for relapse prevention in text-based online environments such as discussion forums. This paper investigates whether language use within a peer-to-peer discussion forum can predict future relapse among individuals treated for AUD. 104 AUD sufferers who had completed residential treatment participated in a mobile phone-based relapse-prevention program, where they communicated via an online forum over the course of a year. We extracted patterns of language use on the forum within the first four months on study using LIWC, a dictionary-based text analysis program. Participants reported their incidence of risky drinking via a survey at four, eight, and twelve months. A logistic regression model was built to predict the likelihood that individuals would engage in risky drinking within a year based on their language use, while controlling for baseline characteristics and rates of utilizing the mobile system. Results show that all baseline characteristics and system use factors explained just 13% of variance in relapse, whereas a small number of linguistic cues, including swearing and cognitive mechanism words, accounted for an additional 32% of the total 45% of variance in relapse explained by the model. Effective models for predicting relapse are needed. Messages exchanged on AUD forums could provide an unobtrusive and cost-effective window into the future health outcomes of AUD sufferers, and their psychological underpinnings. As online communication expands, models that leverage user-submitted text toward predicting relapse will be increasingly scalable and actionable.

CONTACT Rachel Kornfeld rkornfield@wisc.edu School of Journalism and Mass Communication, University of Wisconsin-Madison, 5115 Vilas Hall, 821 University Avenue, Madison, WI 53706.

Alcohol use disorder (AUD) is a serious health condition affecting approximately 33 million Americans in their lifetime (Grant, Goldstein, Saha, & et al., 2015). AUD treatment is challenging, with subsequent relapse being extremely common (Becker, 2008). Studies show that approximately 50% of individuals return to problematic alcohol use within a year following treatment (Moos & Moos, 2006). However, new approaches to treatment that address the full continuum of care, from detoxification to maintenance, hold promise to improve these outcomes (Dennis & Scott, 2007). One such approach, with demonstrated effectiveness in relapse-prevention, involves the use of the Internet and mobile technologies (Quanbeck, Chih, Isham, Johnson, & Gustafson, 2014). Technology-delivered relapse-prevention systems can work, in part, by putting sufferers directly in contact with one another, thus facilitating real-time seeking and sharing of information and social support (Gustafson, McTavish, Chih, & et al., 2014).

In the present paper, we argue that the benefits of online relapse-prevention systems may extend beyond the provision of recovery support. Specifically, we test the possibility that the messages posted by AUD sufferers in online discussion forums could constitute a viable tool for predicting relapse risk. It is now well-established that individuals' language use provides a window into thoughts, feelings and likely actions (Tausczik & Pennebaker, 2010). Since similar ideas can be conveyed in multiple ways, it is not only the explicit content of language, but also its style, that can reveal meaningful psychological differences. Building on this line of work, we propose that it will be possible to identify AUD sufferers at high risk for relapse based on what they say, and how they say it, in online communication forums.

Identifying individuals with elevated risk of relapse is important because it makes it possible to target intensive resources directly at those who most need them. For instance, contact from counselors may not be feasible on a routine basis, but could be justified when focused on high-risk individuals. However, identifying those at highest relapse risk is challenging (Pedersen & Hesse, 2009). Studies have identified a number of risk factors, including severity of alcohol dependence, treatment history, and experience of cravings and negative affect (Witkiewitz, 2011), but assessing such factors after treatment has historically required administering repeated surveys, which is not always possible and can impose a significant burden on participants.

Our approach to assessing relapse risk leverages the text-based nature of online support forums. Drawing on a robust body of research that links individuals' everyday language use with underlying psychological dynamics (see Chung & Pennebaker, 2007, for a review), we propose that it is possible to identify those most likely to relapse based on their patterns of language use as assessed by previously established dictionaries. Specifically, we analyze language use in a mobile phone-based relapse-prevention system via a computational technique, the Linguistic Inquiry and Word Count (LIWC) software (Pennebaker, Booth, & Francis, 2007).

We propose to make several contributions to the literature. While research in health communication has identified linguistic correlates of mental health conditions such as depression and personality disorders (Arntz, Hawke, Bamelis, Spinhoven, & Molendijk,

2012; De Choudhury, Counts, & Horvitz, 2013), there is limited understanding of how psychological traits and states characteristic of AUD recovery might manifest themselves through language. Therefore, we first develop a series of theoretical propositions linking particular categories of language use to AUD relapse risk. Second, we capture language use alongside a range of other potential predictors of relapse assessed through baseline surveys and system use logfiles. Therefore, we not only test the relationships between language and relapse, but we also evaluate unique variance explained by language when accounting for the contributions of other potential predictors. Finally, the applied significance of this study lies in establishing the feasibility of using language captured "in vivo" in online environments to prospectively assess relapse risk. Over half of Americans have downloaded an app to manage their health (Krebs & Duncan, 2015), and mobile technologies are now being adapted to provide ongoing relapse-prevention support (Quanbeck et al., 2014). Therefore, our approach offers potential for massive scalability.

Language Use and Psychological Dynamics

Pennebaker and colleagues' linguistic paradigm proposes that individuals' psychological states and traits manifest themselves in the ways they talk or write (Pennebaker, Mehl & Niederhoffer, 2003). In other words, linguistic choices reflect who we are and what we are doing, thinking, and feeling. This paradigm has been, to a large extent, method-driven: With the development of computerized language analysis techniques, researchers have been able to efficiently extract the linguistic features of texts, and then to correlate these features with psychological processes. Specifically, LIWC (Pennebaker et al., 2007) is a dictionary-based software that counts the frequency of words in a text that belong to theoretically-derived linguistic categories, including parts of speech (e.g., articles, prepositions), psychological processes (e.g., affect, cognition), and personal concerns (e.g., work, religion).

LIWC has been successfully applied in hundreds of studies, yielding a robust body of evidence on the linguistic correlates of psychological processes. Patterns of language use have been shown to validly indicate speakers' personality (Pennebaker & King, 1999), gender (Newman, Groom, Handelman, & Pennebaker, 2008), deceptiveness (Toma & Hancock, 2012), depression (Rude, Gortner, & Pennebaker, 2004), and stress reactions to events (Cohn, Mehl, & Pennebaker, 2004).

According to Pennebaker's paradigm, the linguistic correlates of psychological processes come in two flavors: those describing the content of communication, and those describing the style of communication. That is, we may infer psychological dynamics both from *what* people say and *how* they say it – especially the latter (see Chung & Pennebaker, 2007). Consider the following sentences, which convey the same meaning but in different styles: "Maybe we can work it out" and "This situation is solvable." The first sentence is more inclusive through the use of the "we" pronoun and tentative through the use of the adverb "maybe," whereas the second sentence is less personal through the absence of pronouns and more formal through the use of long words such as "solvable." LIWC attends not only to the manifest content, but also to function words, such as pronouns, that hold sentences together. These words are produced with little conscious awareness and are therefore difficult to

control. Thus, they are especially useful in revealing speakers' underlying psychological dynamics.

Much of the literature tying language use to health outcomes derives from experimental private writing contexts, with participants asked to engage in "expressive writing" accessible only to the experimenters (Pennebaker & Chung, 2011). However, emerging evidence suggests that writing within online forums can also predict well-being (Schwartz et al., 2016). Given the remoteness of the audience and asynchronicity of communication, online forums may replicate aspects of private writing, allowing individuals to candidly relay their thoughts without excessive self-presentation concern (Shim, Cappella, & Han, 2011). Online forums also introduce a social context that may influence how individuals express themselves, and how benefits accrue from their expressions (Liu et al., 2017). Prior work examining online forums for substance abuse recovery suggests that participants use forums to work through personal experiences as well as to build bonds and exchange support (e.g., Coulson, 2014). These uses should generate a range of language relevant to understanding individuals' recovery trajectories.

Linguistic Correlates of Relapse among Individuals in Recovery for AUD

We contend that relapsing among those with AUD is associated with a distinct set of psychological processes and that these processes should be evident in AUD sufferers' online language use, both in content and style. Research suggests that AUD sufferers who are likely to relapse experience particular emotions, thinking patterns, and social and personal concerns that distinguish them from those unlikely to relapse Below, we explain how these psychological processes might manifest themselves through language.

Affect—One strong indicator that AUD sufferers may relapse is the experience of intense negative affect, including depressed mood and acute negative emotional responses to challenges such as cravings or conflicts (Khantzian, 1997; Swendsen et al., 2000), with some in recovery resorting to substance use to soothe distress. At the linguistic level, negative affect can be expected to manifest itself through increased usage of negative affect words (e.g., "sad," "anxious," "upset"). Simply put, individuals who experience negative affect because they are depressed (De Choudhury et al., 2013) or struggling to cope with chronic conditions (Robbins, Mehl, Smith, & Weihs, 2013) tend to reveal that affect through the language they use.

Additionally, the affective experiences of AUD sufferers can include impulsivity, frustration, and a proneness to spontaneous emotion-laden outbursts. Linguistically, one indicator of such impulsivity and frustration would be increased use of swear words (e.g., "damn," "fuck"). Consistent with this reasoning, studies show that individuals increase their production of swear words when depressed (Robbins et al., 2011; Rodriguez, Holleran, & Mehl, 2010) and when struggling to cope with illnesses (Chancellor, Mitra, & De Choudhury, 2016).

In sum, we expect that AUD sufferers who are likely to relapse should produce more negative emotion and swear words than those who are not.

Cognition—Those who struggle in recovery efforts can also find themselves mired in maladaptive thinking patterns. As such, many therapies and mutual help programs center on increasing sufferers' thoughtfulness and ability to generate new insights. For instance, a successful recovery may involve challenging initial perspectives about drinking (Miller, 1983). Individuals with high levels of cognitive processing may therefore stand a higher chance of overcoming alcohol dependence. At the linguistic level, depth of thinking and processing is illustrated by the use of cognitive mechanism words (e.g., "cause," "know," "ought"), which assert relationships between concepts or objects, differentiate between what belongs to a category and what does not, and express knowledge and insight (Tausczik & Pennebaker, 2010). The production of cognitive mechanism words has been associated with adjustment after traumatic events (Pennebaker & Seagal, 1999), coping with cancer (Shaw, Hawkins, McTavish, Pingree, & Gustafson, 2006), and recovering from anorexia (Lyons, Mehl, & Pennebaker, 2006). Because individuals at risk for relapse may engage in less cognitive processing, we expect them to generate fewer cognitive mechanism words.

Furthermore, AUD recovery is broadly characterized by successfully inhibiting impulses to drink. It follows, then, that LIWC's "inhibition" words (e.g. "block," "stop"), which represent a sub-type of cognitive processing focused on control (Pennebaker et al., 2003), should also be produced less by those who will relapse. The use of inhibition words has been linked to successfully managing and resolving conflict in prior research (Brett et al., 2012).

Social Integration—Characteristics of AUD sufferers' social environments also impact relapse risk, with supportive relationships being key in the recovery process. Specifically, those who go on to relapse tend to have smaller, less supportive social networks (Hunter-Reel, McCrady, & Hildebrandt, 2009) and may feel more lonely (Åkerlind & Hörnquist, 1992). The psychological importance of connectedness can manifest itself linguistically through pronoun use. An elevated rate of first-person plural pronouns (e.g., "we") may indicate that individuals see themselves as embedded in a network of social ties. Conversely, an elevated rate of first-person singular pronouns (e.g., "I") may indicate that individuals do not see themselves as socially integrated. Illustrating the importance of social integration in times of distress, research shows that "we" pronouns can predict couples' adjustment to breast cancer (Robbins et al., 2013), and that "I" pronouns" can predict distress and ill health, including suicidality (Brancu, Jobes, Wagner, Greene, & Fratto, 2015), social anxiety (Anderson, Goldin, Kurita, & Gross, 2008), and depression (Rude et al., 2004). Accordingly, we predict that social disconnection, as demonstrated by lower "we" pronoun use and higher "I" pronoun use, predicts relapse.

Furthermore, the substantial distress, stigma, and uncertainty involved in AUD create needs for caring interactions, where individuals feel loved and validated (White, Kelly, & Roth, 2012). The exchange of love and empathetic support has been shown to buffer against psychological stressors in chronic disease management (Yoo et al. 2014). As LIWC 2007 does not include dictionaries corresponding to such bonding, we programmed an additional LIWC dictionary capturing references to love ("love," "loved," loves"). We predict that those likely to relapse lack socially validating interactions, leading them to produce fewer "love" words.

Personal concerns—Finally, addiction often involves a reprioritization of everyday activities, with those struggling in their recovery showing less concern for health (Downey, Rosengren, & Donovan, 2001) and spirituality (Sussman et al., 2013) and exhibiting lower motivation to achieve recovery goals (Aharonovich, Amrhein, Bisaga, Nunes, & Hasin, 2008). Conversely those on their path to recovery demonstrate the opposite pattern – elevated concern with health and spirituality, and higher and more consistent achievement orientation. The salience of these concerns should be illustrated in the content of AUD sufferers' posts.

In the area of health, we expect AUD sufferers who are likely to relapse to ignore the dire consequences of excessive drinking on their health and eventual mortality by decreasing word production in the "health" and "death" LIWC categories. Prior research has found that a decrease in "health" words predicted sobriety struggles among users of online recovery forums (MacLean et al., 2015), and that a decrease in "death" words was associated with the active promotion of anorexia nervosa in online forums (Lyons et al., 2006). That is, mentally disengaging from health- and death-related topics had a negative relationship with adaptive behaviors, whereas talk of health and death seemingly had a protective component.

In the area of spirituality, research shows that recovery often involves the application of religious principles, and many individuals reporting experiences of spiritual "awakening" (Swora, 2004). Religious language has been shown to predict effective coping in the case of cancer patients (Shaw et al., 2007). Thus, increased production of words in the "religion" category in LIWC should be associated with successful recovery, whereas a diminished attention to religious and spiritual themes should be predictive of relapse.

Lastly, in the area of achievement motivation, research shows that setting and pursuing goals represents an important component of successful recovery (McIntosh & McKeganey, 2000). Thus, individuals with low achievement motivation should be more vulnerable to relapse. At the linguistic level, this should manifest in the decreased production of "achievement" words.

In summary, consistent with a multi-dimensional conception of well-being in recovery, we predict relationships between relapse and a range of linguistic features representing affect, social integration, cognitive processing, and personal concerns.

Methods

Study Context

This study examines data from a clinical trial of a smartphone application for AUD relapse-prevention (Gustafson et al., 2014), focusing on the 170 patients who received a smartphone with the Addiction–Comprehensive Health Enhancement Support System (A-CHESS). This application was designed to reduce problem drinking by offering those with AUD access to a variety of informational, communicative, and interactive services including addiction-related content, weekly surveys, and a peer-to-peer discussion forum.

The discussion forum enabled participants to communicate with one another by starting and responding to threads, or simply reading each other's messages. The discussion forum had the following characteristics: 1) it was gender segregated, with messages posted by women visible only to other women, and messages posted by men visible only to other men; 2) it was pseudo-anonymous, as participants used pseudonyms but could upload photos or share real names if they wished; 3) it was moderated by staff who answered queries, directed participants to resources, or started threads to increase engagement; and 4) participation was rolling, such that newly enrolled participants could interact with others who had been in the study for months.

Participants

Participants were at least 18 years old and met criteria for DSM-IV alcohol dependence. They were enrolled in the study after completing residential AUD treatment at one of two treatment organizations in the US, one located in the Midwest and one in the Northeast. The Midwestern treatment program involved cognitive-behavioral therapy, motivational interviewing, and psychoeducation, conducted almost entirely in group therapy, whereas the Northeastern treatment program involved cognitive-behavioral therapy, motivational interviewing, psychoeducation, case management, individual and group therapy, and attendance of community Alcoholics Anonymous meetings. We account for this difference by study site in our analysis. Individuals were invited in the clinical trial only if they did *not* have a history of suicidality or significant developmental, cognitive, or vision impairments that might limit their ability to use the mobile application. Participants who were randomly assigned to the treatment condition were given smartphones loaded with A-CHESS. They provided informed consent and were trained by research coordinators on how to use the mobile application, including the discussion forum. The Institutional Review Boards at the study sites approved the study.

Only those who posted at least one message during the first four months of their participation (N=124) and who were not lost to follow up at twelve months (N=104) were included in the present analysis. The resulting sample had a mean age of 38 years old (SD = 10), was mostly male (57%), and included 83% Caucasians, 13% Blacks, 4% Hispanic/Latinos, and 3% Native American or Alaska Natives (participants could select multiple race/ethnicity categories). About 42% had at least some college education, and about 20% were employed. More than half (62%) reported a history of abusing drugs beyond alcohol (e.g., cocaine, opiates, or methamphetamines). Participants reported baseline abstinence from alcohol of about 65 days.

Measures

Survey measures—The outcome variable (relapse) and several covariates were self-reported through surveys administered at baseline (in person) and at four, eight, and twelve-month follow-ups (by telephone). Relapse was operationalized as engaging in risky drinking within the past month as reported on any follow-up survey. In turn, risky drinking was defined as consumption of more than three standard drinks for women, or more than four standard drinks for men within two hours.

Self-reported covariates collected at baseline included demographics (gender, age), treatment site, days since drinking, and number of prior attempts to quit drinking. Psychological variables were also assessed: participants' predicted likelihood of remaining sober was reported on a scale from one to 100 (a measure of self-efficacy), and the experience of significant relationship problems was a binary response (a measure of relational stability). Both concepts have been identified as meaningful in recovery, with lower self-efficacy and relational stability predicting relapse (Beattie & Longabaugh, 1997; Young, Oei, & Crook, 1991).

Linguistic measures—A total of 2,590 messages were posted on the discussion board during the study. Participants tended to contribute more frequently earlier in the study, with 63% of messages posted during their first four months of participation. We excluded all messages posted after the 4-month survey as well as those posted by forum moderators, resulting in a total corpus of 1,625 messages. All the messages produced by each participant (i.e., those starting new threads and replying to existing threads) were compiled into a single text file. The participants' text files were subjected to analyses via the LIWC software.

LIWC operates by comparing each word in a text to internal dictionaries representing 76 linguistic categories (e.g., pronouns, negative emotion, religion), and then assigning that word to the appropriate categories. Note that a word can belong to multiple categories. For instance, "worried" is both an adjective and a negative emotion word. For each participant, LIWC produced an output specifying the percentage of the total words posted by a participant that belonged to each linguistic category of interest. Therefore, the independent variables in this study were the hypothesized linguistic categories, expressed as a percentage of the total words posted by each participant on the discussion forum.

System log measures—The mobile application system automatically recorded additional covariates that might be related to relapse: (1) the percent of days participants logged onto the system; and (2) the total number of messages posted by each participant. Like the linguistic measures, these covariates were also assessed over the first 4 months on study.

Results

The average number of messages posted by each participant during the first four months of the study was 15.3. Of these, about 34% started new threads, with the rest being replies. Participants posted an average of 410 words to the discussion board (approximately 27 words per message). The LIWC 2007 program recognized about 81% of the words posted. Please refer to Table 1 for descriptive statistics.[1]

Hypotheses were tested using hierarchical logistic regression, which enabled us to assess the extent to which participants' language use on the forum could predict their likelihood of relapse above and beyond baseline characteristics and system use measures. Relapse was entered as a dichotomous dependent variable; self-reported baseline characteristics (days

[1]For correlations between linguistic features, please contact the corresponding author.

since drinking, study site, age, gender, prior quit attempts, relationship problems, self-efficacy) were entered in Step 1; system use measures (days the application was used, total messages posted) were entered in Step 2; and finally the hypothesized linguistic measures were entered in Step 3. The model fit the data well ($X^2 = 62.3$, $p < 0.01$) and explained a total of 45.1% of the variance in relapse. The first block (demographics) accounted for 11.7% of this variance, the second block (system use) for 1.6%, and the linguistic features for a full 31.8%, according to the Cox and Snell R^2 measure.

When it comes to affect, both negative affect indicators (negative emotion and swear words) were significantly associated with increased relapse risk. When it comes to cognition, cognitive mechanism words were significantly associated with decreased relapse whereas inhibition words were significantly associated with increased relapse. For social orientation, neither singular nor plural personal pronoun use was associated with relapse. However, love words were significantly associated with increased relapse. For personal concerns, health and religious words were not significantly associated with relapse. Death words approached significance, in the expected direction of reduced relapse. Increases in achievement words were also associated with expected decreases in the likelihood of relapse. Please refer to Table 2 for regression coefficients and p-values for all the predictors included in the final model.

In summary, our results revealed significant associations to relapse for affect and cognition, and mixed findings for social orientation and personal concerns. The full model, including linguistic cues and covariates, was applied to our full dataset, and achieved 81.7% accuracy in predicting observed relapse outcomes, with accuracy being the proportion of participants correctly classified, including true positives (those predicted to relapse, and observed to relapse) and true negatives (those predicted not to relapse, and observed not to relapse) with a cut-off value of 0.5. The model performed better in classifying non-relapse, with 58 out of the 67 participants who did not relapse accurately classified (86.6%), than relapse, with 27 out of the 37 participants who eventually relapsed accurately classified (73.0%).

Discussion

This study investigated whether it is possible to predict AUD sufferers' likelihood of relapse within one year after treatment based on the language they used in an online discussion forum dedicated to recovery support. Results are encouraging: As few as seven linguistic categories (cognitive mechanism-, inhibition-, negative emotion-, swearing-, love-, death-, and achievement-related words), illustrating AUD sufferers' affect, cognition, social integration and personal concerns, accounted for almost a third of the variance in their relapse likelihood. These language cues predicted relapse above and beyond participants' baseline characteristics (e.g., demographics) and rates of accessing the mobile system, which had limited utility predicting relapse. Our final model including linguistic cues had an accuracy rate of over 80% in identifying AUD sufferers who would relapse within one year, and this was achieved without expensive and demanding methods such as post-treatment surveys.

These findings are consistent with Pennebaker and colleagues' paradigm, according to which psychological processes are evident in what we say (or write). An ample body of literature shows that, by capturing patterns in word use, linguistic coding can reliably predict a range of psychosocial constructs, including writers' personality dimensions, social status, and interaction goals (Chung & Pennebaker, 2007). Some recent work has also applied linguistic analysis to addiction: Collins and colleagues (2009) showed that individuals' language use in open-ended responses could reveal engagement with relapse-prevention principles and concepts from mindfulness-based therapy. Prior research using a hand-coding approach showed that concepts from motivational interviewing manifested in participants' language use during therapy, including "change talk" wherein participants express intention or commitment to change health behavior, with such language predicting treatment outcomes (Amrhein, Miller, Yahne, Palmer, & Fulcher, 2003; Sarpavaara & Koski-Jännes, 2013). In addition, two recent studies leveraged language from social media platforms to predict recovery status as inferred from writers' online sobriety claims (recovery "badges" or mentions of sober time) (MacLean et al., 2015; Tamersoy, De Choudhury, & Chau, 2015). These latter studies support the potential of language features, such as negative affect words, to predict relapse or recovery in a social media context, but are limited by the lack of participant information and systematic outcome measurement.

Our findings are generally aligned with the psychology of AUD relapse. First, they highlight the centrality of affective experience (and its manifestation in language use) to relapse risk. The "self-medication hypothesis" of addiction holds that individuals engage in problematic substance use in order to blunt unpleasant negative affective states (Swendsen et al., 2000), and it would follow that negative affect would often precede binge drinking. Our data is consistent with this claim: language use illustrating affective states, particularly swearing, played an important role in predicting relapse. Some individuals in recovery may lack the self-awareness to use explicitly emotional language (Hopwood, Schade, Matusiewicz, Daughters, & Lejuez, 2015), but swearing potentially allows us to capture their affect indirectly.

Second, we found that cognitive mechanism words (e.g., cause, know) were associated with improved outcomes (i.e., no relapse). Prior studies have suggested that these words index writers' depth of processing of their experiences, which may benefit well-being by allowing writers to adopt new, more functional ways of thinking (Pennebaker & Seagal, 1999). This is consistent with understanding recovery as a "cognitive restructuring" process, where individuals evolve in their perspectives to ultimately change behavior (Steigerwald & Stone, 1999).

We had also proposed that cognition focused on inhibition would be supportive of recovery, but surprisingly, greater use of inhibition language was associated with higher relapse risk. In other words, the more people used inhibition-related words, the less they were able to inhibit their drinking impulses. Given our inability to account for negation with our dictionary-based approach, perhaps this finding reflects discussion of self-control failure, or pledges to enhance self-control in recognition of elevated risk. For instance, the inhibition category would capture messages reading both, "I can stop drinking" and "I can't stop drinking."

We also found mixed performance of language related to social integration. Prior studies using LIWC to predict health events have focused on singular personal pronouns (e.g., “I”) as a gauge of dysfunctional self-focus, but this study showed that rates of using these pronouns did not predict relapse. Further, we failed to replicate prior findings showing that high rates of plural first-person pronouns (e.g., “we”) predict psychological benefits, signifying writers’ embeddedness in social relationships. These findings are surprising given ample evidence that social integration is central to recovery (Hunter-Reel et al., 2009), and that those in recovery take on a collective identity as “alcoholics” or “addicts” (Cain, 1991). It may be that “I” and “we” capture multiple meanings. “I” could indicate isolation but also beneficial self-focus needed to process experiences or take responsibility. “We” could reflect integration in a recovery community (e.g., “We have come too far to give up”), but could also indicate ties to less supportive persons, including substance users (e.g., “We were together for 16 years and I'm not over her”). Multiple meanings could limit the predictive utility of these words.

Finally, we predicted that “love” words would be negatively associated with relapse risk, as they would indicate involvement in caring relationships, yet we found the opposite – the higher the rate of “love” words participants used, the higher their likelihood of relapse. It is possible that expression of love on the forum was relatively shallow, or that those who were struggling with their sobriety and who needed the extra support sought to precipitate bonding as a strategy to overcome their elevated risk. In other words, the use of “love” words may have signified a *bid* for social support, rather than an acknowledgement that this support exists.

We had mixed findings for personal concerns. Despite their prior utility predicting health outcomes in online forums, religion and health words did not predict relapse in this study. Consistent with our hypotheses, our results indicated that achievement words did predict lower relapse risk, and that death words approached significance in the same direction. This latter finding may reflect that those in the study had occasion to reflect on literal instances of death: either their own death, risked through continued substance use (e.g., “I don't wanna die….”), or deaths of those around them (e.g., “yesterday I found out a dear friend of mine in recovery died at a very young age”). Those who identify the seriousness of alcohol use through death talk may be motivated to change their behavior, consistent with literature showing that depicting negative health effects or eliciting fear can motivate behavior change (Peters, Ruiter, & Kok, 2013).

Limitations and Future Directions

Taken together, these findings suggest an opportunity to improve relapse prediction through analyzing text submitted to online peer-to-peer discussion forums. Content analysis of user-submitted messages can inform patient monitoring without increasing burden on participants through surveys or writing assignments, instead leveraging data that participants submit directly as a voluntary component of their self-management efforts. To the extent that models successfully identify at-risk patients, intervention elements can be tailored toward timely relapse prevention or management (e.g., Chih et al., 2014). For instance, we anticipate that predictive models leveraging linguistic features could guide optimization

through matching participants with services and content tailored to stable or unstable recovery trajectories. More intensive treatments or even counselor contact could be offered to those at highest risk. By identifying when intensive treatment is *not* needed, tailoring could also reduce burden and cost.

Our findings are noteworthy given recent efforts to leverage technology to support addiction recovery, including through online peer-to-peer forums (e.g., Coulson, 2014), and mobile-based applications that allow participants to communicate with one another or with counselors (e.g., Gustafson et al., 2014). Whereas stigma can limit discussion of addiction, specialized forums offer an opportunity to comfortably and anonymously raise issues among those who share and understand this experience (Green-Hamann et al., 2011). Mobile capabilities also allow users to post in real-time and from any location as they confront dynamic risk factors such as negative affect, cravings, and stress. For these reasons, language within dedicated forums should be highly relevant to recovery processes and risk assessment.

As online support forums expand their reach, it will be important to distinguish between possible mechanisms by which self-expression brings benefits in these settings. Language use might sometimes *reflect* writers' underlying psychological states, but in other cases it might *affect* psychological changes as writers reflect on their experiences or summon helpful responses from peers. While our study is not designed to distinguish between these accounts, we hope future research will undertake this task. Better understanding the causal mechanisms behind language use and health outcomes would have important applications, including allowing those in the addiction treatment field to recommend writing within particular online venues as a component of recovery. While benefits of self-expression in recovery have been established in clinical trials of private writing (e.g., Meshberg-Cohen, Svikis, & McMahon, 2014), future work might compare effects of private writing to writing in online forums, and the mechanisms by which writing brings benefits in these settings.

As this suggests, future research might also assess how treatment paradigms operate through distinct linguistic mechanisms. For instance, mutual help programs emphasize the value of speaking openly and from personal experience (Kornfield, 2014), and mutual help-influenced forums could operate like private writing in emphasizing expression of personal challenges and facilitating cognitive processing. In contrast, perhaps social integration indicators would have a disproportionate role in predicting outcomes for those in family or couples therapy for AUD.

Our findings should be interpreted in relation to the benefits and shortcomings of computer-aided content analysis, and LIWC in particular. LIWC is useful in part because of its efficient and reliable application to large data sets, its capacity to recognize patterns of language use that would evade detection by human observers, and its validity as established through numerous studies (Tausczik & Pennebaker, 2010). Furthermore, by counting the percent of words (versus percent of messages), it may capture more variance in individuals' language use than manual coding approaches. On the other hand, LIWC 2007 only examines single words or word stems and does not take into account contextual meaning of words or

negation, as noted above. Review by a human coder can be necessary to understand the nature of captured content.

Furthermore, LIWC utilizes pre-defined dictionaries not tailored to addiction. In this study, for instance, LIWC couldn't capture "love" words, so we programmed this category ourselves. Categories such as "health," which showed no significant relationship with relapse, may be too general. Perhaps more sophisticated approaches to natural language processing, such as those leveraging word combinations could capture additional content relevant to relapse, including instances of affirming commitment to recovery or recounting self-redemption, both of which have been predictive of sustained recovery (Amrhein et al., 2003; Dunlop & Tracy, 2013).

This study has additional limitations. First, while individuals were randomized to receive the relapse-prevention mobile application, they were not required to write on the discussion forum, and relapse risk could not be predicted for those who did not write. However, demographic characteristics of those who did and did not write were similar.[2] Second, we did not assess if outcomes reflected whether participants *read* messages containing various content. Reading others' messages could have aided in recovery, or changed how individuals themselves wrote on the forum. Third, it is also important to note that relapse is a common occurrence in AUD recovery, and many individuals relapse multiple times before effectively curbing drinking. While we treated relapse as a binary outcome, there are multiple clinically relevant recovery trajectories, including recurrent relapse. Future work could take a profile analysis approach, comparing language use patterns across a range of recovery trajectories. In addition, future work might monitor outcomes over a longer timeframe to clarify which predictors are suited for assessing short-term versus long-term risk. There may be changes over time in the use of different language features, and these changes might reveal individuals' likely recovery trajectories (i.e., growth curve analysis).

A number of further considerations relate to the applicability of the identified model to new datasets. The participants we examined here were more homogenous than AUD sufferers more broadly since all had completed treatment and were abstinent at baseline. Further, they had completed treatment through one of two treatment organizations. It is possible that language use may correspond to the constructions of successful recovery particular to these organizations and the treatment paradigms employed, including cognitive behavioral therapy, motivational interviewing, and psychoeducation. It is as yet unclear how language use will predict recovery in different forums and among heterogeneous groups of substance users, such as those who have not completed treatment, who use substances other than alcohol, or who subscribe to different views of recovery, perhaps emphasizing moderation over abstinence goals.

[2]For demographic characteristics of participants who did and did not write on the discussion forum, please contact the corresponding author.

Conclusions

Increasingly, seeking and providing support for relapse-prevention occurs in online environments and through mobile technologies. Communication in these environments may not only have an impact on relapse-prevention through increasing mutual support but, as this paper demonstrates, it could also have secondary relapse-prevention benefit through capturing natural language use data from which to predict and respond to individuals' risk level. We found that whereas baseline characteristics and overall volume of activity on the system did not contribute substantially to relapse prediction, incorporating linguistic features into our model offered major improvements. Future studies can build on this work by assessing language models in new AUD populations, incorporating growth curve models for language features, and refining language coding techniques in order to assess other recovery-relevant language elements.

References

Aharonovich E, Amrhein PC, Bisaga A, Nunes EV, Hasin DS. Cognition, commitment language, and behavioral change among cocaine-dependent patients. Psychology of Addictive Behaviors. 2008; 22:557–562. DOI: 10.1037/a0012971 [PubMed: 19071981]

Åkerlind I, Hörnquist JO. Loneliness and alcohol abuse: A review of evidences of an interplay. Social Science & Medicine. 1992; 34:405–414. DOI: 10.1016/0277-9536(92)90300-F [PubMed: 1566121]

Amrhein PC, Miller WR, Yahne CE, Palmer M, Fulcher L. Client commitment language during motivational interviewing predicts drug use outcomes. Journal of Consulting and Clinical Psychology. 2003; 71:862–878. DOI: 10.1037/0022-006X.71.5.862 [PubMed: 14516235]

Anderson B, Goldin PR, Kurita K, Gross JJ. Self-representation in social anxiety disorder: Linguistic analysis of autobiographical narratives. Behaviour Research and Therapy. 2008; 46:1119–1125. DOI: 10.1016/j.brat.2008.07.001 [PubMed: 18722589]

Arntz A, Hawke LD, Bamelis L, Spinhoven P, Molendijk ML. Changes in natural language use as an indicator of psychotherapeutic change in personality disorders. Behaviour Research and Therapy. 2012; 50:191–202. DOI: 10.1016/j.brat.2011.12.007 [PubMed: 22317755]

Beattie MC, Longabaugh R. Interpersonal factors and post-treatment drinking and subjective wellbeing. Addiction. 1997; 92:1507–1521. DOI: 10.1111/j.1360-0443.1997.tb02871.x [PubMed: 9519493]

Becker HC. Alcohol dependence, withdrawal, and relapse. Alcohol Research and Health. 2008; 31:348–361. DOI: 10.1016/B978-0-12-405941-2.00019-5 [PubMed: 23584009]

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. Archives of Suicide Research. 2016; 20:438–450. DOI: 10.1080/13811118.2015.1040935 [PubMed: 26219609]

Brett JM, Olekalns M, Friedman R, Goates N, Anderson C, Lisco CC. Sticks and stones: Language, face, and online dispute resolution. Academy of Management Journal. 2007; 50:85–99.

Cain C. Personal stories: Identity acquisition and self-understanding in Alcoholics Anonymous. Ethos. 1991; 19:210–253. DOI: 10.1525/eth.1991.19.2.02a00040

ChancellorS, , MitraT, , De ChoudhuryM. Proceedings of the 2016 CHI Conference on Human Factors in Computing SystemsSanta Clara, CA: ACM; 2016Recovery amid pro-anorexia: Analysis of recovery in social media; 21112123

Chih M-Y, Patton T, McTavish FM, Isham AJ, Judkins-Fisher CL, Atwood AK, Gustafson DH. Predictive modeling of addiction lapses in a mobile health application. Journal of Substance Abuse Treatment. 2014; 46:29–35. DOI: 10.1016/j.jsat.2013.08.004 [PubMed: 24035143]

ChungC, , PennebakerJ. The psychological functions of function words. In: FiedlerK, editorSocial communicationNew York, NY: Psychology Press; 2007343359

Cohn MA, Mehl MR, Pennebaker JW. Linguistic markers of psychological change surrounding September 11, 2001. Psychological Science. 2004; 15:687–693. DOI: 10.1111/j.0956-7976.2004.00741.x [PubMed: 15447640]

Collins SE, Chawla N, Hsu SH, Grow J, Otto JM, Marlatt GA. Language-based measures of mindfulness: Initial validity and clinical utility. Psychology of Addictive Behaviors. 2009; 23:743–749. DOI: 10.1037/a0017579 [PubMed: 20025383]

Coulson NS. Sharing, supporting and sobriety: A qualitative analysis of messages posted to alcohol-related online discussion forums in the United Kingdom. Journal of Substance Use. 2014; 19:176–180. DOI: 10.3109/14659891.2013.765516

De ChoudhuryM, , CountsS, , HorvitzE. Proceedings of the 5th Annual ACM Web Science ConferenceNew York, NY, USA: ACM; 2013Social media as a measurement tool of depression in populations; 4756

Dennis M, Scott CK. Managing addiction as a chronic condition. Addiction Science & Clinical Practice. 2007; 4:45–55. [PubMed: 18292710]

Downey L, Rosengren DB, Donovan DM. Sources of motivation for abstinence: A replication analysis of the Reasons for Quitting Questionnaire. Addictive Behaviors. 2001; 26:79–89. DOI: 10.1016/S0306-4603(00)00090-3 [PubMed: 11196294]

Dunlop WL, Tracy JL. Sobering stories: Narratives of self-redemption predict behavioral change and improved health among recovering alcoholics. Journal of Personality and Social Psychology. 2013; 104:576.doi: 10.1037/a0031185 [PubMed: 23437925]

Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, Hasin DS. Epidemiology of DSM-5 alcohol use disorder: Results from the national epidemiologic survey on alcohol and related conditions iii. Journal of the American Medical Association Psychiatry. 2015; 72:757–766. DOI: 10.1001/jamapsychiatry.2015.0584 [PubMed: 26039070]

Green-Hamann S, Campbell Eichhorn K, Sherblom JC. An exploration of why people participate in Second Life social support groups. Journal of Computer-Mediated Communication. 2011; 16:465–491. DOI: 10.1111/j.1083-6101.2011.01543.x

Gustafson DH, McTavish FM, Chih MY, Atwood AK, Johnson RA, Boyle MG, Isham A. A smartphone application to support recovery from alcoholism: A randomized clinical trial. Journal of the American Medical Association Psychiatry. 2014; 71:566–572. DOI: 10.1001/jamapsychiatry.2013.4642 [PubMed: 24671165]

Hopwood CJ, Schade N, Matusiewicz A, Daughters SB, Lejuez CW. Emotion regulation promotes persistence in a residential substance abuse treatment. Substance Use & Misuse. 2015; 50:251–256. DOI: 10.3109/10826084.2014.977393 [PubMed: 25404467]

Hunter-Reel D, McCrady B, Hildebrandt T. Emphasizing interpersonal factors: An extension of the Witkiewitz and Marlatt relapse model. Addiction. 2009; 104:1281–1290. DOI: 10.1111/j.1360-0443.2009.02611.x [PubMed: 19549057]

Khantzian EJ. The self-medication hypothesis of substance use disorders: A reconsideration and recent applications. Harvard Review of Psychiatry. 1997; 4:231–244. DOI: 10.3109/10673229709030550 [PubMed: 9385000]

Kornfield R. (Re)working the program: Gender and openness in Alcoholics Anonymous. Ethos. 2014; 4:415–439. DOI: 10.1111/etho.12062

Krebs P, Duncan DT. Health app use among US mobile phone owners: A national survey. JMIR mHealth and uHealth. 2015; 3:e101.doi: 10.2196/mhealth.4924 [PubMed: 26537656]

Liu Y, Kornfield R, Shaw B, Shah DV, McTavish F, Gustafson DH. When support is needed: Social support solicitation and provision in an online Alcohol Use Disorder forum. Digital Health. 2017; 3:1–16. DOI: 10.1177/2055207617704274

Lyons EJ, Mehl MR, Pennebaker JW. Pro-anorexics and recovering anorexics differ in their linguistic Internet self-presentation. Journal of Psychosomatic Research. 2006; 60:253–256. DOI: 10.1016/j.jpsychores.2005.07.017 [PubMed: 16516656]

MacLeanD, , GuptaS, , LembkeA, , ManningC, , HeerJ. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social ComputingVancouver, BC, Canada: ACM; 2015 Feb. Forum77: An analysis of an online health forum dedicated to addiction recovery; 15111526

McIntosh J, McKeganey N. Addicts' narratives of recovery from drug use: Constructing a non-addict identity. Social Science & Medicine. 2000; 50:1501–1510. DOI: 10.1016/S0277-9536(99)00409-8 [PubMed: 10741584]

Meshberg-Cohen S, Svikis D, McMahon TJ. Expressive writing as a therapeutic process for drug-dependent women. Substance Abuse. 2014; 35:80–88. DOI: 10.1080/08897077.2013.805181 [PubMed: 24588298]

Miller WR. Motivational interviewing with problem drinkers. Behavioural Psychotherapy. 1983; 11:147–172.

Moos RH, Moos BS. Rates and predictors of relapse after natural and treated remission from alcohol use disorders. Addiction. 2006; 101:212–222. DOI: 10.1111/j.1360-0443.2006.01310.x [PubMed: 16445550]

Newman ML, Groom CJ, Handelman LD, Pennebaker JW. Gender differences in language use: An analysis of 14,000 text samples. Discourse Processes. 2008; 45:211–236.

Pedersen MU, Hesse M. A simple risk scoring system for prediction of relapse after inpatient alcohol treatment. The American Journal on Addictions. 2009; 18:488–493. DOI: 10.3109/10550490903205983 [PubMed: 19874170]

PennebakerJW, , BoothRJ, , FrancisME. Linguistic inquiry and word count: LIWCAustin, TX: LIWC.net; 2007

PennebakerJW, , ChungCK. Expressive writing: Connections to physical and mental health. In: FriendmanHS, editorThe Oxford handbook of health psychologyNew York, NY: Oxford University Press; 2011417437

Pennebaker JW, King LA. Linguistic styles: Language use as an individual difference. Journal of Personality and Social Psychology. 1999; 77:1296–1312. DOI: 10.1037/0022-3514.77.6.1296 [PubMed: 10626371]

Pennebaker JW, Mehl MR, Niederhoffer KG. Psychological aspects of natural language. use: Our words, our selves. Annual Review of Psychology. 2003; 54:547–577. DOI: 10.1146/annurev.psych.54.101601.145041

Pennebaker JW, Seagal JD. Forming a story: The health benefits of narrative. Journal of Clinical Psychology. 1999; 55:1243–1254. DOI: 10.1002/(SICI)1097-4679(199910)55:10<1243::AID-JCLP6>3.0.CO;2-N [PubMed: 11045774]

Peters G-JY, Ruiter RAC, Kok G. Threatening communication: A critical re-analysis and a revised meta-analytic test of fear appeal theory. Health Psychology Review. 2013; 7(Suppl. 1):s8–s31. DOI: 10.1080/17437199.2012.703527 [PubMed: 23772231]

Quanbeck A, Chih M-Y, Isham A, Johnson R, Gustafson D. Mobile delivery of treatment for Alcohol Use Disorders. Alcohol Research: Current Reviews. 2014; 36:111–122. [PubMed: 26259005]

Robbins ML, Focella ES, Kasle S, López AM, Weihs KL, Mehl MR. Naturalistically observed swearing, emotional support, and depressive symptoms in women coping with illness. Health Psychology. 2011; 30:789–792. DOI: 10.1037/a0023431 [PubMed: 21574707]

Robbins ML, Mehl MR, Smith HL, Weihs KL. Linguistic indicators of patient, couple, and family adjustment following breast cancer. Psycho-Oncology. 2013; 22:1501–1508. DOI: 10.1002/pon.3161 [PubMed: 22887054]

Rodriguez AJ, Holleran SE, Mehl MR. Reading between the lines: The lay assessment of subclinical depression from written self-descriptions. Journal of Personality. 2010; 78:575–598. DOI: 10.1111/j.1467-6494.2010.00627.x [PubMed: 20433631]

Rude S, Gortner E-M, Pennebaker J. Language use of depressed and depression-vulnerable college students. Cognition and Emotion. 2004; 18:1121–1133. DOI: 10.1080/02699930441000030

Sarpavaara H, Koski-Jännes A. Change as a journey—Clients' metaphoric change talk as an outcome predictor in initial motivational sessions with probationers. Qualitative Research in Psychology. 2013; 10:86–101. DOI: 10.1080/14780887.2011.586256

SchwartzHA, , SapM, , KernML, , EichstaedtJC, , KapelnerA, , AgrawalM, , KosinskiM. Proceedings of the Pacific Symposium on BiocomputingKohala Coast, Hawaii: World Scientific; 2016Predicting individual well-being through the language of social media; 516527

Shaw B, Han JY, Kim E, Gustafson D, Hawkins R, Cleary J, Lumpkins C. Effects of prayer and religious expression within computer support groups on women with breast cancer. Psycho-Oncology. 2007; 16:676–687. DOI: 10.1002/pon.1129 [PubMed: 17131348]

Shaw BR, Hawkins R, McTavish F, Pingree S, Gustafson DH. Effects of insightful disclosure within computer mediated support groups on women with breast cancer. Health Communication. 2006; 19:133–142. DOI: 10.1207/s15327027hc1902_5 [PubMed: 16548704]

Shim M, Cappella JN, Han JY. How does insightful and emotional disclosure bring potential health benefits? Study based on online support groups for women with breast cancer. Journal of Communication. 2011; 61:432–454. DOI: 10.1111/j.1460-2466.2011.01555.x [PubMed: 25568496]

Steigerwald F, Stone D. Cognitive restructuring and the 12-step program of Alcoholics Anonymous. Journal of Substance Abuse Treatment. 1999; 16:321–327. DOI: 10.1016/S0740-5472(98)00052-X [PubMed: 10349605]

Sussman S, Milam J, Arpawong TE, Tsai J, Black DS, Wills TA. Spirituality in addictions treatment: Wisdom to know… what it is. Substance Use & Misuse. 2013; 48:1203–1217. DOI: 10.3109/10826084.2013.800343 [PubMed: 24041182]

Swendsen JD, Tennen H, Carney MA, Affleck G, Willard A, Hromi A. Mood and alcohol consumption: An experience sampling test of the self-medication hypothesis. Journal of Abnormal Psychology. 2000; 109:198–204. DOI: 10.1037/0021-843X.109.2.198 [PubMed: 10895557]

Swora MG. The rhetoric of transformation in the healing of alcoholism: The twelve steps of Alcoholics Anonymous. Mental Health, Religion & Culture. 2004; 7:187–209.

TamersoyA, , De ChoudhuryM, , ChauDH. Proceedings of the 26th ACM Conference on Hypertext & Social MediaAnkara, Turkey: ACM; 2015 Aug. Characterizing smoking and drinking abstinence from social media; 139148

Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology. 2010; 29:24–54. DOI: 10.1177/0261927X09351676

Toma CL, Hancock JT. What lies beneath: The linguistic traces of deception in online dating profiles. Journal of Communication. 2012; 62:78–97. DOI: 10.1111/j.1460-2466.2011.01619.x

White WL, Kelly JF, Roth JD. New addiction-recovery support institutions: Mobilizing support beyond professional addiction treatment and recovery mutual aid. Journal of Groups in Addiction & Recovery. 2012; 7:297–317. DOI: 10.1080/1556035X.2012.705719

Witkiewitz K. Predictors of heavy drinking during and following treatment. Psychology of Addictive Behaviors. 2011; 25:426–438. DOI: 10.1037/a0022889 [PubMed: 21480681]

Yoo W, Namkoong K, Choi M, Shah DV, Tsang S, Hong Y, Gustafson DH. Giving and receiving emotional support online: Communication competence as a moderator of psychosocial benefits for women with breast cancer. Computers in Human Behavior. 2014; 30:13–22. DOI: 10.1016/j.chb.2013.07.024 [PubMed: 24058261]

Young RM, Oei TPS, Crook GM. Development of a drinking self-efficacy questionnaire. Journal of Psychopathology and Behavioral Assessment. 1991; 13:1–15. DOI: 10.1007/BF00960735

Table 1

Linguistic features of AUD discussion board messages.

	Number LIWC words	Examples	Mean	SD
Word Count	--	--	410.45	543.06
Dictionary Words	~4,500	--	80.98	14.32
1st person singular	12	I, me, mine	7.28	4.13
1st person plural	12	We, our, us	0.39	0.72
Cognitive mechanisms	730	Cause, know, ought	13.82	5.95
Inhibition	111	Block, constrain, stop	0.83	1.30
Negative emotions	499	Hurt, ugly, nasty	1.24	1.27
Swearing	53	Damn, piss, fuck	0.09	0.16
Love	3	Love, loved, loves	0.29	1.26
Health	236	Clinic, flu, pill	0.64	0.98
Religion	159	Altar, church, mosque	0.60	0.10
Death	62	Bury, coffin, kill	0.05	0.13
Achievement	186	Earn, hero, win	2.39	2.06

Note. Number LIWC words refers to the number of words per category in the LIWC dictionary. Mean refers to the mean percentage of words in each category present on the discussion board. Word count is expressed as a raw number.

Table 2

Summary of final logistic regression model predicting relapse by month 12.

Variable	*B*	SE *B*	*p*
Baseline characteristics			
Gender	1.752	0.872	0.044
Age	−0.057	0.036	0.115
Study Site	−1.087	0.878	0.216
Sober days	−0.015	0.009	0.098
Prior Quit Attempts	−0.046	0.039	0.233
Relationship Problems	1.794	1.051	0.088
Sobriety chance	−0.019	0.022	0.381
System use characteristics			
Days used	0.014	0.018	0.432
Messages posted	−0.042	0.028	0.138
Linguistic features			
1st person singular	0.007	0.085	0.939
1st person plural	0.940	0.733	0.199
Cognitive mechanisms	−0.166	0.077	0.031
Inhibition	1.283	0.488	0.009
Negative emotions	0.806	0.346	0.020
Swearing	9.275	2.947	0.002
Love	1.503	0.482	0.005
Health	−0.119	0.350	0.734
Religion	0.030	0.325	0.972
Death	−6.211	3.299	0.060
Achievement	−1.248	0.469	0.008

Note. For gender, 0=male and 1=female. For relationship problems, 0=no and 1=yes.