Proceedings of the Seventeenth International AAAI Conference on Web and Social Media (ICWSM 2023)

The Pursuit of Peer Support for Opioid Use Recovery on Reddit

Duilio Balsamo,[1,2] Paolo Bajardi,[3,4] Gianmarco De Francisci Morales,[3,4]
Corrado Monti,[3,4] Rossano Schifanella[2,4]
[1] Bocconi University, Milan, Italy
[2] University of Turin, Turin, Italy
[3] CENTAI, Turin, Italy
[4] ISI Foundation, Turin, Italy
duilio.balsamo@unibocconi.it, {paolo.bajardi, gdfm, corrado.monti}@centai.eu, rossano.schifanella@unito.it

Abstract

Individuals suffering from Opioid Use Disorder and other socially stigmatized conditions often rely on peer support groups to find comfort and motivation while treating their condition. Many may face barriers in accessing peer support treatment, such as shame and social stigma, seclusion, or mobility restrictions. In this study, we quantitatively characterize the potential of the Reddit community in offering these individuals an online alternative to receiving peer support. By analyzing the social interactions of thousands of users during the start of opioid use recovery, we uncover that a particular Reddit community exhibits many characteristics similar to in-person peer support groups, featuring the exchange of support, trust, status, and similar experiences. We find that the supportive behavior of this community nudges users to change their personal behavior, and promotes abandoning opioid-related communities in favor of recovery-oriented relationships. Finally, we find that recognition, acknowledgment, and knowledge exchange are the most relevant factors in sustained engagement with the recovery community. Given this evidence, we suggest that this online community may constitute a complement or a surrogate to peer support groups when in-person meetings are not desirable or possible. Our work might inspire harm reduction policies and interventions to favor successful rehabilitation and is fundamental for future research about the use of digital media for recovery support.

Introduction

In the last two decades, the unprecedented growth of deaths due to opioid drugs has overwhelmed the United States (CDC 2019), and in 2020 an estimated 2.7 million people aged 12 or older suffered from an Opioid Use Disorder (OUD) (SAHMSA 2021). During and after a formal recovery treatment, individuals suffering from OUD are usually encouraged to participate in *peer support groups*, whereby individuals with similar conditions share mutual non-professional assistance (Tracy et al. 2016; Strang et al. 2020). Compared to professional clinical advice, *peer support* has the possibility of reaching a deeper level of trust, understanding, and acceptance due to the communal background and shared experience of the individuals (Mead and MacNeil 2006). Peer support groups have been proven effective in treating substance use and conditions related to mental health (Reif et al. 2014; Bassuk et al. 2016). Studies reveal that people who participate in peer support groups benefit from reduced substance use and relapse rates, while incentivizing other peer users in engagement with treatment (De Choudhury and De 2014; Sharma et al. 2020). Key elements of the experience in such groups are *social support*, i.e., positive psycho-social interactions with others with whom there is mutual *trust* and concern (Sarason et al. 1983), experiential *knowledge*, and peer mentoring ”through the shared experience of emotional and psychological pain” (Mead, Hilton, and Curtis 2001). These groups can be especially beneficial during the initial stages of recovery, as they provide encouragement to endure, status through community reinforcement (Brown, Tang, and Hollman 2014), and empowerment by promoting self-confidence in participants (Barak, Boniel-Nissim, and Suler 2008; Reif et al. 2014). Although initial involvement in a peer support group is crucial to trigger new life-changing habits, the reasons for staying are usually to maintain social support and develop self-esteem (Schutt and Rogers 2009). Moreover, participation and active engagement in peer support groups is a key predictor of achieving recovery (Donovan and Wells 2007; Best and Lubman 2012) and of sustaining recovery (Etheridge et al. 1999). Consequently, individuals with higher attendance levels have shown statistically significant improvements over time in self-esteem, self-efficacy, and community activism-autonomy (Vayshenker et al. 2016).

Alongside the physical struggle posed by drug detoxification, such as withdrawal symptoms, a further challenge to successful substance use recovery is managing the uncertainties of a drastic lifestyle change (Tracy et al. 2016; Bassuk et al. 2016). Existing literature suggests that the mechanism of recovery is a process of *social group change* and *social identity change* (Best et al. 2016). In the process, a person’s identity shifts from being characterized by the membership to a group whose norms and values revolve around substance use to being defined by the membership to a new group whose norms and values encourage recovery. Hence, a crucial step for a successful remission and sustained engagement with supportive relationships of recovery (Haslam et al. 2016; McKeganey 2000) is a “public renegotiation” that starts with a public announcement of the intention to change lifestyle (Stall and Biernacki 1986).

Individuals suffering from OUD and other socially stigmatized conditions (Betton et al. 2015) such as mental health issues (De Choudhury and De 2014), sex- or gender-related issues (Nobles et al. 2018; Saha et al. 2019), and physical health issues (Enes et al. 2018) have now the alternative to engage with online communities spontaneously. People in need of peer support are increasingly finding a safe place to share their experiences in online social media platforms that provide pseudonymity, such as Reddit (Bunting et al. 2021; Andalibi et al. 2016; Sowles et al. 2018).

Reddit is a social content aggregation website where users can post, comment, and vote on content, and which hosts several user-created, topical communities called *subreddits*. In some specific communities, users freely discuss substance use. Conversely, individuals in substance use recovery can refer to several specialized subreddits, such as the *r/OpiatesRecovery* community, which is the subject of this study.

These communities have been previously used in research studies to track the spatial patterns of opioid interest in the US (Balsamo, Bajardi, and Panisson 2019), to investigate alternative substances used for opioid use recovery (Chancellor et al. 2019), to study the use of non-medical routes of administration for opioid drugs (Balsamo et al. 2021), and to understand the types and sources of stigma surrounding Opioid Use Treatment (Kepner, Meacham, and Nobles 2022). Eshleman et al. (2017) used machine learning to predict the propensity of starting recovery and Yang et al. (2019) of relapsing. However, no previous large-scale study has investigated *whether it is possible to find the social characteristics of peer support groups in this community*. Answering this question requires analyzing the type of social relationships that form in this group and studying how they impact the behavior of users on the platform. D'Agostino et al. (2017) provided qualitative evidence on this phenomenon thanks to the analysis of 100 comments, by identifying OUD standard diagnostic criteria (DSM-5) (APA 2021) in such community and suggesting its therapeutic potential.

In this work, we quantitatively answer this question by implementing a purposely-built computational pipeline with natural language processing and statistical analysis. Considering publicly available Reddit data from 2015 to 2019, we analyze thousands of posts by pseudonymous Reddit authors in recovery, and we identify more than 2k who explicitly disclose the timeline of their recovery process on the platform.

We uncover the presence of peer support dynamics in their interactions with *r/OpiatesRecovery*, such as the exchange of *Support*, *Trust*, *Status*, and *Similarity* (Choi et al. 2020). This measurable social dynamic suggests that Reddit users who begin opioid use recovery consider this community a safe place to find support. Moreover, we measure how peer support evolves as participants progress in their recovery, from two months before the start of the recovery up to two months after, by using Interrupted Time Series (ITS) analysis. This way, we find that *r/OpiatesRecovery* also fosters *social group change* and *behavioral change*, including the abandonment of other opioid-related communities.

Lastly, we highlight which types of online social interactions impact the most on the engagement and attachment of the users to the recovery community. We find that recognition and acknowledgment from peers are crucial factors, as well as the exchange of *Knowledge* and *Support*. Our results show that participation in the *r/OpiatesRecovery* community presents many characteristics proper of peer support treatment. As such, we suggest that this community should be carefully considered by public health authorities and health professionals, e.g., as a venue to establish first contact with individuals in need of support, to complement existing in-person programs, or as a viable surrogate when participation in physical meetings is undesirable or impossible.

Methods

The digital cohort subject of this study is a group of Reddit authors that publicly disclose the beginning of their recovery process from opioid use. Using machine learning, we identify the posts in which such authors share their achievements during the recovery process on *r/OpiatesRecovery*. Thanks to these posts, we estimate the date t_0 in which $N = 2125$ Reddit authors start their opioid use recovery process. Then, we align their posting timelines according to the respective t_0, and we study the authors' behavior from two months before to two months after the start of the recovery.

Estimating the Start of Recovery

This section describes how we build our dataset of *r/OpiatesRecovery* participants, where each recovering author is paired with an estimated day t_0. The pipeline is composed of three stages: (i) we collect a sample of submissions where we manually identify whether the author self-reports the time elapsed since beginning recovery; (ii) we use this annotated dataset to train a machine learning model to find submissions of this type among all the available ones; (iii) we use regular expressions to estimate the starting date of the recovery process for each recovering author. Let us now describe in detail each of these stages.

Dataset and annotation. We collect the data from the Pushshift Reddit repository (Baumgartner et al. 2020). This publicly available dataset consists of the content of all the public subreddits published on the platform since 2007 (Medvedev, Lambiotte, and Delvenne 2018). In particular, we analyze the textual part of the submissions and the comments collected from 2015 to 2019, with more than 265k posts. To create a small hand-curated dataset to build our model, we collect a random sample of 1000 user submissions on *r/OpiatesRecovery*. We manually check the submission's title for the presence of self-reports of recovery that include references to the time elapsed since its beginning. Specifically, we annotate as positive examples the posts referring to personal and firsthand experiences of recovery that also contain clear temporal markers which indicate the time spent in recovery, e.g., "Today I'm two weeks clean". We annotate all the other posts as negative examples, including those that refer to the detoxification of others, to relapses, or those related to other subjects. This annotation process produces 223 positive posts and 777 negative ones. We split this set into two stratified datasets for training and validation, which respectively contain 70% and 30% of the examples while preserving the positive-negative ratio.

Temporal expressions of recovery. We use this labeled dataset to train and test a machine-learning model capable of identifying submissions that contain self-reported recovery periods. We test the classification performance of 4 well-established machine-learning models for text classification (Logistic Regression, Random Forest, Decision Tree, and Support Vector Machine, implemented in *scikit-learn*).[1] We perform model selection on the training dataset with 5-fold cross-validation and grid-search for hyper-parameter tuning. Given the slight class imbalance in the training and validation sets, we choose the ROC-AUC score as the target performance metric. As a text preprocessing step we remove punctuation and perform lemmatization with *NLTK*.[2] We also automatically transform the text snippets referring to numeral quantities into digits using the *text2digits* library.[3] Then, we transform the preprocessed corpus of texts into a numerical vector representation of word frequencies using a bag-of-words approach with optional *tf-idf* weighting. Optionally, prior to the text vectorization step, we mask the words referring to known time expressions with a unique time-expression token to allow the model for more generalization. To do so, we develop a series of regular expression rules to match expressions referencing a certain number of hours, days, weeks, months, and years, such as *'Day 5', '2 months'*. We also account for complex forms which might include unrelated terms, e.g., *'1 painful week'*. Based on the best validation performance of AUC $= 0.942$ and Matthew's correlation coefficient (Boughorbel, Jarray, and El-Anbari 2017) MCC $= 0.847$ we select the pipeline consisting of time-expression masking, tf-idf weighting, and logistic regression model as the most suitable for our task. After retraining the best model on the entire training set, we use it to predict a score on all the remaining unlabelled submissions posted by users on *r/OpiatesRecovery*. Out of the 18 186 total submissions, 4227 (23%) are predicted as positive by our model, i.e., with a high probability of the presence of recovery declaration and specific mention of time spent in recovery.

Extraction of temporal references. Next, we focus on isolating the textual expressions regarding the time elapsed while in recovery and converting them into their numeric equivalent t_d expressed in *days*. We apply a series of regular expressions rules to capture the numerical part n of textual expressions that reference an elapsed time. We convert each post p in a vector $n_p = [n^h, n^d, n^w, n^m, n^y]_p$ which contains the number of *hours*, *days*, *weeks*, *months*, and *years* expressed. Then, we transform each component into the corresponding equivalent in days. For simplicity, we standardize the duration of months to 30 days and that of years to 365 days. We proceed by assigning a potential *recovery time* t_d in days to each post. For posts containing only one time expression, we output the sum of the components of n_p in days.

[1]https://scikit-learn.org

[2]https://www.nltk.org

[3]https://pypi.org/project/text2digits

In case a post contains multiple time expressions, we identify the following heuristic: if the time expressions are tied by conjunction, e.g., "1 week and 4 days" we sum the two vectors. In case of multiple expressions separated by 'in', 'after', 'from', or 'for', we consider only the first time expression. In all other cases, we discard the submission. Our procedure finds 3805 submissions which contain an expression of recovery time t_d, belonging to $N = 2125$ distinct Reddit users (26% of the total users who created at least one submission on *r/OpiatesRecovery*).

Estimate of recovery time. To estimate a candidate date t_0 when the author of a post started the recovery process, we subtract t_d days from the creation date of the submission. Since each recovering author might have written multiple submissions containing recovery progress, either referring to the same recovery period or multiple separate ones, we associate a set of candidate recovery starting dates to each author, one for each submission. The estimation of the t_0 from a submission can suffer from errors introduced by our pipeline and unclear reporting. Therefore, we rely on DBSCAN (Ester et al. 1996), a well-established density-based clustering procedure, to identify consistent reports and to discard outliers. In this paradigm, a cluster identified by the algorithm for a recovering author represents a set of reports consistently pointing to a temporal neighborhood around the same t_0, and outliers represent spurious or incorrect reports. Those t_0 reported only once but not conflicting with other consistent periods are accepted. Finally, we select the most frequent t_0 in each cluster as the representative recovery starting date.

The time from t_0 to the posting date of the last submission associated with the same cluster identifies an uninterrupted period of abstinence from substance use. Hence, when recovery periods overlap, we discard the period with fewer reports, which is more uncertain. Since OUD is clinically considered a chronic disease, those affected might slip back into relapse periods. We find this behavior also in our cohort, where 302 out of 2125 authors report multiple non-overlapping recovery periods, separated by periods where relapse may have occurred. We consider only the first recovery period reported by an author in case of multiple ones.

Behavioral Fingerprinting

To create a fingerprint of the behavior of the recovering authors in the platform, we use a series of daily measures y_t. The measures rely either on the content or on the metadata of the posts on a given day and are described in detail below. Similarly to Fan et al. (2019) and ElSherief et al. (2021), we align individual timelines according to a point-wise event and aggregate the metrics of the aligned timelines to observe temporally-resolved population-level measures. This procedure allows us to measure the behavioral change of the recovering authors over time and the evolution of their social interactions. Specifically, we collect the posting activity of each recovering author in a time window of 60 days before and after the respective starts of recovery t_0, including all the comments the community wrote in reply to those posts. We remove all the posts in which the recovering authors disclosed their elapsed time in recovery so as not to introduce biases due to data collection. Finally, we align the authors' timelines by offsetting each one to its respective start of recovery t_0, i.e., we express each timeline in a range of $t \in [-60, 60]$ days. Based on existing literature, we consider two different types of measurements, which reflect two aspects of the behavior of users.

Social features. The first category of measurements considers the types of interactions happening among the recovering authors and the community. These features reflect our aim to measure changes in social feedback and peer support experienced during recovery. We rely on a set of Long short-term memory (LSTM) classifiers by Choi et al. (2020), pre-trained on a corpus of Reddit posts, to determine the binary presence in the posts of ten basic *social dimensions of conversation and relationships* (Deri et al. 2018). These include *Support, Trust, Status, Similarity, Fun, Conflict, Knowledge, Power, Romance* and *Identity*. Table 1 reports a brief description and some anonymized examples of the social dimensions present in our dataset. Given a textual input, each of these recurrent neural network classifiers outputs a score between 0 and 1 that reflects the presence of one of the said social dimensions. The classifiers perform better on medium-sized text, so we split each submission and comment into sentences, and we assign a score for every dimension of conversation to each sentence. For each dimension of conversation, we max-pool the scores of the sentences to compute the final score relative to the submission or comment (Monti et al. 2022). Then, we transform the scores to determine the binary presence of a said social dimension in each post. Rather than estimating the optimal thresholds for each classifier, we adopt an approach that mitigates the intrinsic biases of the classifiers: we evaluate the quartiles of the distribution of all the scores predicted by the classifier for each specific social dimension and binarize those scores based on their membership to the upper quartile. Given that our goal is to evaluate the changes and the evolution in time in a comparative manner rather than estimating the crude quantity of the social dimensions exchanged, this approach is better suited for our analytical framework. Finally, we report for each author the daily fraction of posts containing such dimension over the total number of daily posts, representing the daily aggregate relative measure of the social dimensions exchanged.

Activity features. The second category of measurements corresponds to activity-related measures, from which we can observe social group change and shifts in personal engagement with the recovery community. These activity-related features rely on collected posts' metadata, which include the score and the subreddit of the posts. We follow a similar procedure to the one adopted for evaluating the social dimensions of conversation: we group the submissions and the comments created or received daily by each recovering author, and we provide daily measures of the activity-related features. In particular, we evaluate the number of submissions created or comments received (*N. Posts*), the number of unique subreddits where the authors have posted (*N. Subreddits*), and the number of distinct authors with whom the recovering authors have interacted (*N. Contacts*). Moreover, we break down their activity on *r/opiates* and *r/OpiatesRecovery* by computing the ratios of submissions and comments on these subreddits w.r.t. the total number of posts of each author in the day (respectively, *Share Opiates* and *Share OpR*). Similarly, we compute the relative share of interactions with unique users on *r/opiates*, on *r/OpiatesRecovery*, and on all the other subreddits combined (respectively, *Share Contacts Opiates*, *Share Contacts OpR* and *Share Contacts Other*). In addition, as measures related to social feedback, we evaluate the sum of the scores assigned to the posts (*Sum. Scores*) and the daily average valence of the posts (*Avg. Valence*), computed via *VADER*.[4] To avoid potential biases caused by different levels of baseline engagement of the users with the platform, e.g., users with a different posting frequency on Reddit, we normalize the measures based on raw counts (e.g., number of posts, number of contacts) by computing z-scores on the entire timeline of each author.

Statistical Assessment of Behavioral Shift

Our dataset consists of numerous cross-sectional observations of metrics of behavior corresponding to multiple authors, spanning two months before and after the beginning of recovery. To quantify the presence of change in the behavior of the recovering authors and the community conditioned to the start of recovery t_0, we apply two different methodologies. Both methods consider all the data points and provide estimates at the population level but aggregate the data differently. The first method, *Average Behavioral Shift*, enables control for a single user's behavior, and the second, *ITS*, for their longitudinal behavior.

Average Behavioral Shift. This measure quantifies how much a given behavior has changed on average at the population level when comparing the recovery period to the one preceding it, controlling for single-user behavior. It first compares the average values of a particular metric computed before and after the start of recovery, separately for each author. Then it provides the typical shift at the population level by averaging the shifts of all the users. Formally, it is defined as:

$$\delta_y = \frac{1}{|U|} \sum_{i \in U} \bar{y}^i_{t \geq t_0} - \bar{y}^i_{t < t_0} \tag{1}$$

where $\bar{y}^i_{t \geq t_0}$ and $\bar{y}^i_{t < t_0}$ are the average values of the measure y_t for user $i \in U$, respectively after and before the beginning of recovery, and U is the set of all recovering authors. By means of $B = 1000$ rounds of bootstrap, shuffling the authors' timelines, we evaluate p-values to check for the statistical significance of each average shift δ_y. Statistical significance is coded as *: $P \leq .05$, **: $P \leq .01$, ***: $P \leq .001$ throughout the paper.

Interrupted Time Series. To understand how the behavior of the individuals who started opioid use recovery changes throughout the recovery process, we employ an Interrupted Time Series (ITS) analysis (McDowall, McCleary, and Bartos 2019). ITS is a quasi-experimental technique that

[4]https://github.com/cjhutto/vaderSentiment

Social dimension	Definition	Example
Support	Giving emotional or practical aid and companionship	Hope you will sort it out soon! – I suggest to all others who are struggling, to post on here daily to help you stay focused.
Trust	Will of relying on the actions or judgments of another	I want to make her proud and I'm not going to let her or myself down. – It's been a good day, feeling that restless legs aren't too bad and I'll have a good night.
Similarity	Shared interests, motivations or outlooks	When I am doing something I feel normal, but when I sit around it really is the only thing on my mind. – Now that is just my own personal experience but I think a shit load of addicts feel the same.
Status	Conferring status, appreciation, gratitude, or admiration	Just wanted to say it to you, good work. – Well done, keep it up!
Power	Having power over behavior and outcomes of another	Been to 7 meetings in the last 5 days. Start work again tomorrow. – Keep up the good work!
Fun	Experiencing leisure, laughter, and joy	If I can pass every class while practically numb, I must be a god while sober. This is hilarious. – My indoor cat makes me happy, he gets pretty stupid whenever he has the opportunity to eat plants.
Conflict	Contrast or diverging views	It hurts me to know that nobody sponsors me. – We generally hide behind an image of being bad-ass when in actuality we can't even tolerate the most minor symptoms from withdrawal.
Knowledge	Exchange of ideas or information; learning, teaching	I was taking small doses (15-22.5mg) of DXM twice a day, I read about the positive results it has on opiate withdrawal in a clinical research trial. – Only used recommended dosage of lope for the constant shits of the first week, didn't try to use it to suppress withdrawal suffering.
Romance	Intimacy among people (sentimental or sexual)	You were a genuinely beautiful soul. – I love you people.
Identity	Shared sense of belonging to the same group	Spent everything to be there as it is the greatest affirmation of her faith, as being a good catholic? – Strength, freedom, autonomy, are all characteristics of people that find sobriety.

Table 1: Social dimensions definition and examples from our dataset.

allows the estimation of the causal effect of an intervention happening at a defined time point in the absence of a counterfactual. Thanks to its inferential powers, ITS has been widely adopted to assess the effects of health and policy interventions (Bernal, Cummins, and Gasparrini 2017; Chandrasekharan et al. 2017; Tian and Chunara 2020). We consider the start of recovery t_0 as the intervention and study its effect on the time series of a given behavior y_t. After aligning the timelines of the users on t_0, for each dimension y_t we fit an ordinary least squares regression model:

$$y_t = \beta_0 + \beta_1 t + \beta_2 D_t + \beta_3 P_t \tag{2}$$

where D_t is a dummy variable with a null value before the start of recovery, and P_t is the progressive number of days in recovery. Considering a measure y_t such as *Support* received from the community, the regression coefficients β_0 and β_1 reflect respectively the pre-recovery level of *Support* exchanged and its trend approaching t_0. β_2 quantifies the immediate effect of the intervention, i.e., the quantity of immediate *Support* received at the start of recovery. Finally, β_3 reflects the trend change in *Support* during recovery.

Prediction of Next-Week Participation

To study the role of peer support on recovery, we set up a binary classification task with the next-week-participation to *r/OpiatesRecovery* as the dependent variable and the social dimensions of conversation exchanged as independent variables. With our framework, we are able to estimate the start of recovery but not quantify its outcome in terms of the success of recovery. For this reason, we perform further analysis on the weekly participation to *r/OpiatesRecovery* since the start of recovery, i.e., a binary variable that indicates the presence of either submissions or comments on this subreddit on a said week, as a proxy of attachment to the community and possibly of progress in recovery. We set up different classification tasks by using the most relevant social ties of conversation exchanged on *r/opiatesrecovery* by each user at week t (either expressed by the authors, by the community, or both) to predict the presence of activity of the same user during the following week $t+1$. To account for different behaviors at different stages of progression in recovery, we include the week of prediction as an independent variable.

Results

Presence and Evolution of Peer Support

Our goal is to understand whether the hallmarks of peer support might be found in the social interactions of the Reddit community and whether these interactions affect the behavior of its users. Figure 1a and Table 2 report the *Average Behavioral Shift* δ_y, i.e., the population-average difference in the exchange of a social dimension y between the periods after and before the start of recovery t_0. The measure uses the fraction of daily posts that contain a specific social dimension in each period. Quantities referring to posts written by the Reddit authors undergoing recovery (*recovering authors* henceforth) are in blue, while those for the ones written by the rest of the community as a response (*community* henceforth) are in red. There is a positive average shift with a strong statistical significance in the exchange of the social

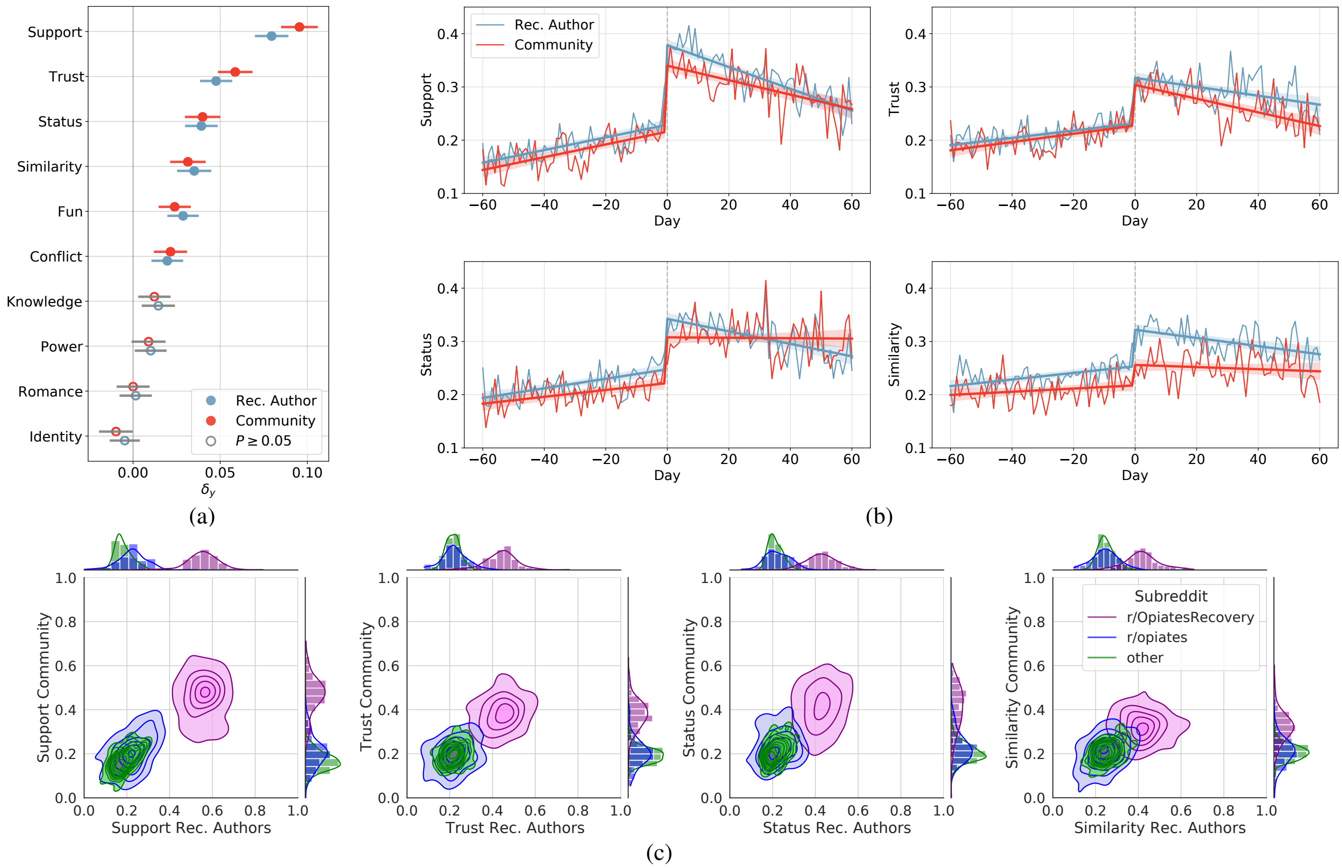


Figure 1: Average behavioral shift (a) and Interrupted time series (b) of the social dimensions of conversation regarding the authors and community, respectively reported in blue and red. In panel (a) dots correspond to the average value $\bar{\delta}_y$ and bars report the Standard Error of the Mean. Values not statistically significant are indicated with hollow markers and gray bars. In (b), the thin lines report the daily average of the studied social dimensions, in a temporal span ranging from two months before to two months after the start of detox t_0. The thick lines represent the corresponding ITS fits, with 95% confidence intervals as shaded areas. The plots in panel (c) show the distribution of daily averages of the share of authors' and community's posts containing certain social dimensions. The plots report the bivariate distributions relative to the three communities studied, shown as kernel density plots, with the marginal authors' and community's distributions on the axes.

dimensions that can be related to peer support, namely *Support*, which increases the most ($8\%^{***}$ on average), followed by *Trust*, *Status*, and *Similarity*. The shifts in the exchange of these social dimensions present matching behaviors across the two directions of conversation, from the recovering authors to the community and vice versa. The shift in the social content the recovering authors receive from the community is even greater. A likely explanation for this phenomenon is that the authors in recovery *shift* their interactions towards the more supportive communities. Finally, we do not find statistical evidence of a shift in the exchange of the other dimensions of *Knowledge*, *Power*, *Romance*, and *Identity*.

ITS analysis confirms these results and provides further details on their temporal unfolding. Figure 1b reports the daily average as well as the corresponding ITS fits of the four most relevant social dimensions of conversation to characterize the evolution of peer support. Both directions of exchange of *Support*, *Trust*, and *Status*, along with the *Similarity* dimension expressed by the authors, exhibit similar behavior: a slight but significant increasing trend approaching t_0, and a strong positive shift at the start of recovery (e.g., $+14.9\%^{***}$ for *Support*). As the recovery progresses, the exchanges of *Support* and *Trust* return to their baseline levels at a faster rate compared to the growth measured before t_0. The ITS analysis finds lower or no significance for the shifts in the exchange of *Conflict*, *Knowledge Romance*, and *Identity*. Table 4 reports values and statistical significance of the coefficients of the ITS fits.

To better characterize the contribution of each social group to the exchange of peer support among users in recovery, we perform a comparative analysis across Reddit communities. We compare the average daily quantity of social dimensions exchanged on three different subreddit types: (i) *r/OpiatesRecovery*; (ii) *r/opiates*, a subreddit predominantly focused on non-medical use of opioids; (iii) the aggregate of all the other subreddits. These measures reflect the typical share of posts in which a given social dimension is expressed in a day on the specific subreddits by the average recovering author or by the respective community. Figure 1c shows the bivariate distributions of the measures relative to the three communities in the study, shown as kernel density plots together with the marginal distributions for authors and interacting community. While the distribution of the average *Support*, *Trust*, *Status*, and *Similarity* exchanged on *r/opiates* and on other subreddits overlap, the distributions relative to *r/OpiatesRecovery* are mostly non-overlapping with the other communities and clearly display a stronger exchange of these social dimensions. For example, an author of *r/OpiatesRecovery* shares, on an average day, an average of 55.3% of posts containing a *Support* message and receives a similar one at about the same rate (46.8%). The typical user on *r/opiates*, conversely, shares and receives on average less than half of the messages expressing *Support* (22.4% and 20.6%, respectively). These results indicate that, indeed, the *r/OpiatesRecovery* community stands out for its peer support characteristics, contrary to the other main community of discussion for opioids and the general Reddit environment.

Community Shift and Social Feedback

Next, we study the behavior of the recovering authors by evaluating several metrics based on their posting activity. The results regarding the average behavioral shift and the ITS analysis (Figure 2, Table 2, and Table 4) are coherent evidence of increased social activity on the platform, of the presence of social feedback, and of social group change, with authors shifting interactions from the opioid-using community towards the recovery one. In particular, Figure 2a shows the average difference δ_y of all the population-level activity measures y relative to the periods after and before the start of recovery. There is a significant decrease in the share of posts created by the recovering authors on *r/opiates* (*Share Opiates* $\delta_y = -0.068^{***}$) with a simultaneous positive shift in those exchanged on *r/OpiatesRecovery* (*Share OpR* $\delta_y = 0.327^{***}$). This migration of authors from the opioid-use-oriented community to the recovery-oriented community indicates *social group change*, a crucial component of the process of social identity change that is part of the recovery process.

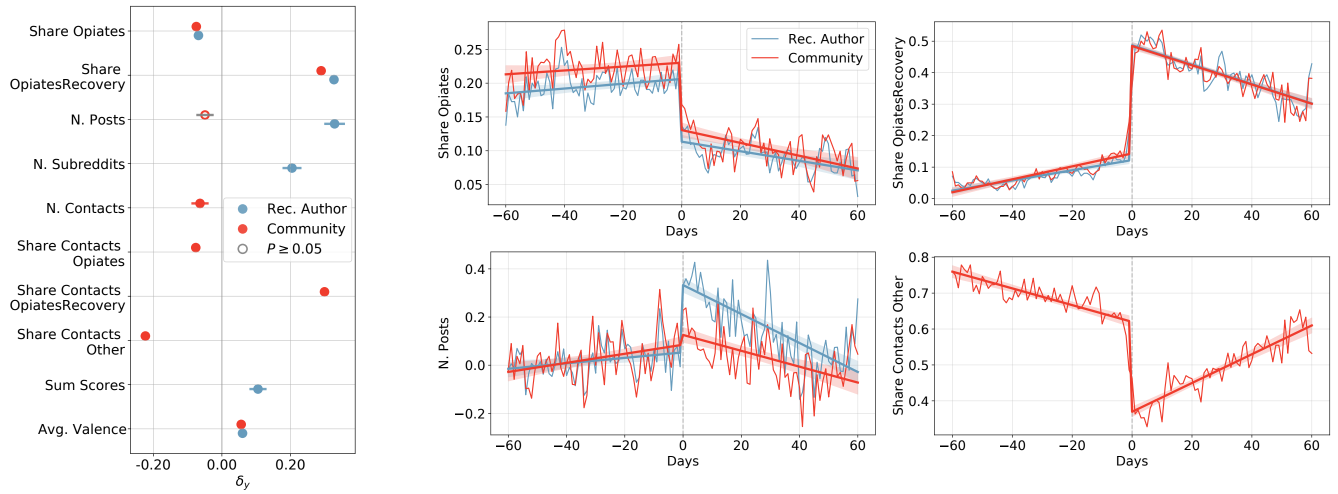
After beginning recovery, despite the positive increase in the number of posts created on average by the recovering authors (*N. Posts* $\delta_y = 0.329^{***}$) and in the number of subreddits used (*N. Subreddits* $\delta_y = 0.205^{***}$), we measure a significant decrease in *N. Contacts* ($\delta_y = -0.064^{*}$), the size of the social group that engages with the posts of the recovering author. Moreover, we observe a shift away from non-opioid-related subreddits and towards *r/OpiatesRecovery* (*Share Contacts Opiates* $\delta_y = -0.076^{***}$, *Share Contacts OpR* $\delta_y = 0.300^{***}$). Lastly, we measure a significant average increase in the score assigned by the community to the posts of the recovering authors (*Sum Scores* $\delta_y = 0.105^{***}$) and in the sentiment (*Avg. Valence* $\delta_y = 0.060^{***}$) of the posts exchanged with the community. These measures indicate positive *social feedback dynamics* taking place between authors and their community during recovery.

The ITS plots in Figure 2b show a significant positive increase at t_0 in the average number of posts created (*N. Posts*). This posting activity decreases as the recovery progresses, and participation in *r/opiates* and *r/OpiatesRecovery* follows opposite temporal evolutions. The daily share of posts on *r/opiates* is steady up to the beginning of the recovery, when it drastically reduces to half the previous participation, with further progressive reduction as the recovery advances. Conversely, the participation in the recovery subreddit (*Share OpR*) grows steadily before the start of detox, and at t_0 its prevalence increases 5-fold, reaching on average more than 40% of the total amount of recovering authors' posts on Reddit. During recovery, while still maintaining a considerable share of activity of around 30% after two months, participation in this subreddit tapers off in favor of activity on other subreddits. Lastly, the interactions with users on subreddits other than *r/opiates* and *r/OpiatesRecovery* (*Share Contacts Other*), after a decreasing trend before recovery and a significant reduction at t_0, progressively increase during recovery, despite a measurable average decrease (*Share Contacts Other* $\delta_y = -0.223^{***}$). These results indicate that the authors who persevere in recovery interact more and more with users of other communities after an initial shrinkage of interactions not driven by opioid consumption/recovery.

Two proxy measures of social feedback, namely *Sum Scores*, the score assigned by the community, and *Avg. Valence*, the valence of the comments received or expressed, increase at t_0 as shown in Figure 4. During recovery, the scores assigned to the authors' posts go back on average to pre-recovery levels and the valence of the comments received by the users stabilizes to consistently higher levels than the valence expressed by their submissions. These findings suggest that the *r/OpiatesRecovery* community is very prompt at giving positive feedback to the users who begin recovery who are also keener to express positive valence.

Sustained Engagement with Recovery Community

Lastly, we investigate how social relationships on *r/OpiatesRecovery* impact authors' engagement in the recovery community, which we consider a proxy of commitment to recovery. Via this analysis, we find that peer recognition and the exchanges of *Support* and *Knowledge* have beneficial effects on the participation of the authors in the recovery community. The panels in Figure 3 show the results of multiple binary regression tasks predicting the participation of the authors to *r/OpiatesRecovery* during the following week. The models use the social dimensions exchanged on the subreddit in the current week. The figure shows the regression coefficients for three different models that consider as covariates: the social dimensions evaluated on posts of the recovering authors (blue), the ones of community posts (red), or the two combined (purple) (see also Table 3). The models considering the features expressed by the authors (purple and blue models) show that *Power*, *Knowledge*, and *Support* are the social dimensions that contribute the most to community attachment among those expressed by the recovering authors. Considering the social dimensions expressed by the community (purple and red models), we find slightly different results regarding the contribution of *Support* and *Power*.



(a)

(b)

Figure 2: Average behavioral shift (a) and Interrupted Time Series (b) of the activity dimensions regarding the recovering authors and community, respectively reported in blue and red. In panel (a) dots correspond to $\bar{\delta_y}$ and bars report the Standard Error of the Mean. Values not statistically significant are indicated with hollow markers and gray bars. In panel (b), the thin lines report the daily average of the studied social dimensions, in a temporal span ranging from two months before to two months after the start of detox t_0. The thick lines represent the corresponding ITS fits, with 95% confidence intervals as shaded areas.

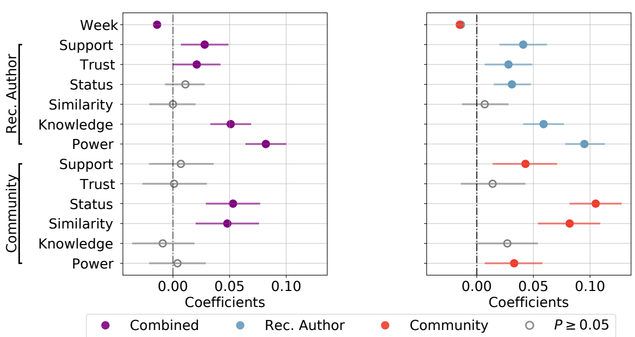


Figure 3: Coefficients and 95% confidence interval of the binary regression tasks predicting authors' participation in *r/OpiatesRecovery* on the next available week. Values not statistically significant are indicated with hollow markers and gray bars. The colors represent the coefficients of three different models, which use as covariates the combination of all the social dimensions (purple) or just the ones evaluated on authors' posts (blue) or Community posts (red).

Nevertheless, these models indicate that *Status* and *Similarity* expressed by the community are the most influential community counterparts to the *Power* and *Support* expressed by the authors. These results also confirm that *acknowledgment* and the *presence of peers* are important factors in community attachment. The first is shown as the well-known *Power-Status* dynamic among users in recovery, and the second is indicated by high *Similarity* exchange. The presence of these social dimensions is not only a defining factor of this community, but it also contributes to keeping its users engaged.

Discussion

Our analysis empirically showed that the *r/OpiatesRecovery* community shares some fundamental traits with peer support groups. Among the social dimensions found in the messages exchanged on the subreddit, *Similarity* resonates with the peer nature of the group. In addition, users exchange *Support*, which conveys emotional, social, or practical help. *Trust* and mutual concern also contribute to characterizing this community as a peer support group. Moreover, our findings on *Status* and *Power* exchanges resonate with two main goals of peer support: providing community reinforcement, empowerment, and fostering self-esteem in participants. Combined, these results confirm the presence of peer-support-like interactions in the community.

Behavior and social group change. Our results show that the authors who undertake recovery significantly increase their social activities, measured as their engagement on the platform, but decrease the size of the social group they interact with. Further analysis of the community where these interactions take place unveils that this decrease is largely due to lower engagement with users participating in *r/opiates*. The interactions with *r/OpiatesRecovery* users increase drastically at the beginning of the recovery, while those on other non-opioid-related subreddits regain relative importance during its progression. These results corroborate the hypothesis of a drastic behavioral shift, with users losing ties with some of their past opioid-related relationships while opening social ties with a restricted but focused group of people in recovery. Thus, the peer support characteristics of the *r/OpiatesRecovery* community attract those users who are in pursuit of help with their recovery. This need drives the change of behavior and the community shift observed in our results, centered around the start of the recovery. From that day, these users also experience social group change, thus shifting away from communities discussing the recreational usage of opioids. Lastly, we find that such users start changing their online behavior even before the start of recovery, slightly increasing their participation in the recovery subreddit in advance.

Barriers to peer support. While *r/OpiatesRecovery* shares many traits with traditional peer support groups, it also presents some key differences. The lower friction in moving inside and outside the community is in fact a double-edged sword. Barriers to traditional peer support group participation involve accessibility and personal factors, including time conflicts, difficulties sharing feelings in person, privacy concerns, and not being familiar with anyone who is a group member (Rapp et al. 2006; Biegel and Song 1995). The modality of access to online peer support groups – written and asynchronous– differs from the in-person experience and allows users to access its content at any time. Thanks to these lower barriers of entry, lack of stigma, and the reassuring presence of peers, the *r/OpiatesRecovery* community offers easy access to peer support, thus possibly reducing the attrition to begin recovery (Vogel, Wade, and Hackler 2007; Wright 2016). Reddit's online and pseudonymous nature may ease the participation of all those who would otherwise suffer social stigma from their social circle (family, colleagues, and acquaintances) due to public admission of OUD (Kepner, Meacham, and Nobles 2022). Moreover, physical restrictions might hinder access to peer support groups, while online groups might offer the alternative to participating in a community with similar characteristics from anywhere. This opportunity can help those who live in secluded places, far from in-person peer support groups, or with mobility restrictions. The recent Covid-19 pandemic has exacerbated this issue by limiting mobility and discouraging group gatherings (Mellis, Potenza, and Hulsey 2021; Krawczyk et al. 2021; Blanco, Compton, and Volkow 2021).

Conversely, the impersonal nature of online interactions could also limit their efficacy. In-person meetings foster deeper relationships among peers, where participants are fully engaged and focused. Instead, *r/OpiatesRecovery*'s online and pseudonymous nature might cause low accountability and easy opt-out. The lack of accountability towards one's social circle may thwart the motivation needed to overcome obstacles in the recovery process. For this reason, we have investigated which factors help retain engagement: receiving peer recognition, support, and knowledge from the community contributes to a sustained engagement with the recovery group. Moreover, our results showed that the active commitment of the recovering authors to beginning recovery and their participation in *r/OpiatesRecovery* enables the community to support them. In fact, we empirically saw that the shift in behavior and community support, engaging with the Reddit community, are in sync with the actual beginning of recovery—even when its disclosure comes after.

Future work. Among the limitations of this work, we acknowledge that the recovering authors selected in our study were not clinically diagnosed with Opioid Use Disorder. Also, we did not investigate the role of peer recovery coaches, which cover the crucial function of leading the discussion in many peer support groups. Future work should focus on addressing these two aspects. Despite recent efforts to enhance the availability of remote treatments programs for substance use disorders through telemedicine (Fiacco, Pearson, and Jordan 2021), virtual meetings (Galanter, White, and Hunter 2021), and specialized health online fora (MacLean et al. 2015), peer support groups still face many challenges in delivering their services remotely. Our work suggests that thanks to its peculiar characteristics, the *r/OpiatesRecovery* represents a valuable example of such a service and should be taken into consideration as a complementary treatment service. If properly advertised, the availability of highly supportive and informative content about opioid use recovery might spur and give conscience to many of those who face OUD and to whom is in doubt about beginning recovery. The combination of online and offline peer support groups has been found to be beneficial to users (Strand, Eng, and Gammon 2020). Therefore, taking our findings as an example, public health authorities might consider creating or developing similar online-based peer support groups to complement classical treatment and peer recovery services. Furthermore, since our work shows that the authors in this subreddit are in pursuit of peer support, policies could be implemented to establish contact with users in particular need, offering them personalized tools to continue their therapy.

Leveraging the richness of information on social media to investigate on otherwise hard-to-reach populations and exploiting the analytical power of natural language processing for understanding at a fine level how users in need interact, we believe that our work sheds light on the importance of open, easily reachable and non-judgemental online tools for support. This work is a first attempt at bringing the scientific community a step closer to understanding which types of communication are suitable in such delicate environments, and might potentially apply to other substance-use-related conditions and mental health issues.

Appendix

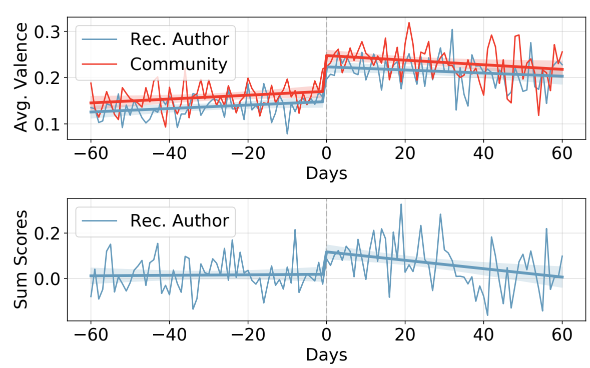


Figure 4: ITS of the Avg. Valence and Sum Scores.

$\times 10^{-1}$		Author δ_y	Author SEM	Community δ_y	Community SEM
Activity Dimensions	Share Op	-0.68^{***}	(0.08)	-0.75^{***}	(0.09)
	Share OpR	3.27^{***}	(0.12)	2.90^{***}	(0.13)
	N. Posts	3.29^{***}	(0.30)	-0.49	(0.25)
	N. Subreddits	2.05^{***}	(0.28)	—	
	N. Contacts	—		-0.64^{*}	(0.25)
	Share cont. Op	—		-0.76^{***}	(0.09)
	Share cont. OpR	—		3.00^{***}	(0.13)
	Share Cont. Other	—		-2.23^{***}	(0.13)
	Sum Scores	1.05^{***}	(0.25)	—	
	Avg. Valence	0.60^{***}	(0.11)	0.56^{***}	(0.11)
Social Dimensions	Support	0.80^{***}	(0.10)	0.95^{***}	(0.11)
	Trust	0.48^{***}	(0.09)	0.40^{***}	(0.10)
	Status	0.39^{***}	(0.09)	0.59^{***}	(0.10)
	Similarity	0.35^{***}	(0.10)	0.21^{*}	(0.10)
	Fun	0.29^{**}	(0.09)	0.24^{**}	(0.09)
	Conflict	0.20^{*}	(0.09)	0	(0.09)
	Knowledge	0.15	(0.10)	-0.10	(0.10)
	Power	0.10	(0.09)	0.31^{**}	(0.10)
	Romance	0.02	(0.09)	0.12	(0.09)
	Identity	-0.05	(0.09)	0.09	(0.10)

Table 2: Average Behavioral Shift and Standard Error of the Mean ($\times 10^{-1}$) . Significance: *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$.

$\times 10^{-2}$		Combined β	Combined 95% CI	Author β	Author 95% CI	Community β	Community 95% CI
	Week	-1.4^{***}	± 0.2	-1.4^{***}	± 0.3	-1.5^{***}	± 0.3
Author	Support	2.8^{*}	± 2.1	4.1^{***}	± 2.1	—	—
	Knowledge	5.1^{***}	± 1.8	5.9^{***}	± 1.8	—	—
	Power	8.2^{***}	± 1.8	9.5^{***}	± 1.8	—	—
	Similarity	-0.0	± 2.1	0.7	± 2.1	—	—
	Status	1.1	± 1.8	3.1^{***}	± 1.7	—	—
	Trust	2.1^{*}	± 2.1	2.8^{**}	± 2.1	—	—
Community	Support	0.7	± 2.8	—	—	4.3^{**}	± 2.8
	Knowledge	-0.9	± 2.7	—	—	2.7	± 2.8
	Power	0.4	± 2.5	—	—	3.3^{*}	± 2.6
	Similarity	4.8^{***}	± 2.8	—	—	8.2^{***}	± 2.8
	Status	5.3^{***}	± 2.4	—	—	10.5^{***}	± 2.3
	Trust	0.1	± 2.8	—	—	1.4	± 2.8
	Adj. R^2	0.201		0.195		0.173	
	AIC	1.103×10^4		1.110×10^4		1.136×10^4	
	BIC	1.113×10^4		1.115×10^4		1.141×10^4	

Table 3: Coefficients ($\times 10^{-2}$) of the next-week participation logistic regression. Significance: *: $P \leq .05$, **: $P \leq .01$, ***: $P \leq .001$

Ethical Statement

This work follows the guidelines and the ethical considerations by Eysenbach and Till (2001); Moreno et al. (2013); Ramírez-Cifuentes et al. (2020). All the results provide aggregated estimates and do not include any information on individuals. The users in our study were fully aware of the public nature and free accessibility of the content they posted since the subreddits are of public domain, are not password-protected, and have thousands of active subscribers. Reddit's pseudonymous accounts make the retrieval of the true identity of users unlikely. Nevertheless, as a further privacy measure, the authors' names were anonymized before using the data for analysis. Therefore, our research did not require informed consent.

$\times 10^{-1}$		β_0	β_1	β_2	β_3
Author	Sh. Op	2.06^{***}	0.0^{*}	-0.93^{***}	-0.01^{***}
	Sh. OpR	1.22^{***}	0.02^{***}	3.65^{***}	-0.05^{***}
	N. Posts	0.53^{***}	0.01^{*}	2.8^{***}	-0.07^{***}
	N. Subreddits	0.57^{***}	0.01	1.2^{***}	-0.03^{***}
	Sum Scores	0.18	0.0	0.99^{***}	-0.02^{**}
	Avg. Valence	1.48^{***}	0.0^{*}	0.75^{***}	-0.01^{*}
Community	Sh. Op	2.3^{***}	0.0	-1.00^{***}	-0.01^{***}
	Sh. OpR	1.43^{***}	0.02^{***}	3.41^{***}	-0.05^{***}
	N. Posts	0.85^{***}	0.02^{***}	0.41	-0.05^{***}
	N. Contacts	0.79^{***}	0.02^{***}	0.19	-0.05^{***}
	Sh. cont. Op	2.29^{***}	0.0	-1.02^{***}	-0.01^{***}
	Sh. cont. OpR	1.51^{***}	0.02^{***}	3.52^{***}	-0.05^{***}
	Sh. cont. Oth.	6.2^{***}	-0.02^{***}	-2.5^{***}	0.06^{***}
	Avg. Valence	1.7^{***}	0.0^{*}	0.77^{***}	-0.01^{**}
Author	Support	2.29^{***}	0.01^{***}	1.49^{***}	-0.03^{***}
	Trust	2.32^{***}	0.01^{***}	0.85^{***}	-0.02^{***}
	Status	2.48^{***}	0.01^{***}	0.95^{***}	-0.02^{***}
	Similarity	2.53^{***}	0.01^{***}	0.68^{***}	-0.01^{***}
	Fun	2.54^{***}	0.0^{*}	0.51^{***}	-0.01^{***}
	Conflict	2.37^{***}	0.0	0.32^{***}	-0.01^{*}
	Knowledge	2.49^{***}	0.0^{**}	0.37^{***}	-0.01^{***}
	Power	2.46^{***}	0.0^{*}	0.26^{***}	-0.0
	Romance	2.72^{***}	0.0	0.19^{**}	-0.0
	Identity	2.32^{***}	0.0	0.14	0.0^{*}
Community	Support	2.16^{***}	0.01^{***}	1.24^{***}	-0.03^{***}
	Trust	2.28^{***}	0.01^{***}	0.76^{***}	-0.02^{***}
	Status	2.22^{***}	0.01^{***}	0.86^{***}	-0.01^{*}
	Similarity	2.17^{***}	0.0	0.38^{***}	-0.0^{*}
	Fun	2.11^{***}	0.0	0.38^{***}	-0.0
	Conflict	2.27^{***}	0.0	-0.03	-0.01^{**}
	Knowledge	2.58^{***}	0.0^{*}	0.12	-0.01^{***}
	Power	2.44^{***}	0.0^{**}	0.51^{***}	-0.02^{***}
	Romance	2.16^{***}	0.0	0.29^{***}	-0.0
	Identity	2.45^{***}	0.0	0.06	-0.0

Table 4: Coefficients of ITS analysis ($\times 10^{-1}$). Significance: * $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$.

References

Andalibi, N.; Haimson, O. L.; De Choudhury, M.; and Forte, A. 2016. Understanding social media disclosures of sexual abuse through the lenses of support seeking and anonymity. In *Proceedings of the 2016 CHI conference on human factors in computing systems*, 3906–3918.

APA. American Psychiatric Association. 2013. Diagnostic and statistical manual of mental disorders: DSM-5. *Arlington, VA*.

Balsamo, D.; Bajardi, P.; and Panisson, A. 2019. Firsthand opiates abuse on social media: monitoring geospatial patterns of interest through a digital cohort. In *The World Wide Web Conference*, 2572–2579.

Balsamo, D.; Bajardi, P.; Salomone, A.; and Schifanella, R. 2021. Patterns of routes of administration and drug tampering for non-medical opioid consumption: data mining and content analysis of reddit discussions. *J. of Medical Internet Research*, 23(1): e21212.

Barak, A.; Boniel-Nissim, M.; and Suler, J. 2008. Fostering empowerment in online support groups. *Computers in human behavior*, 24(5): 1867–1883.

Bassuk, E. L.; Hanson, J.; Greene, R. N.; Richard, M.; and Laudet, A. 2016. Peer-delivered recovery support services for addictions in the United States: A systematic review. *J. of substance abuse treatment*, 63: 1–9.

Baumgartner, J.; Zannettou, S.; Keegan, B.; Squire, M.; and Blackburn, J. 2020. The Pushshift Reddit Dataset. *arXiv preprint arXiv:2001.08435*.

Bernal, J. L.; Cummins, S.; and Gasparrini, A. 2017. Interrupted time series regression for the evaluation of public health interventions: a tutorial. *Int. J. of epidemiology*, 46(1): 348–355.

Best, D.; Beckwith, M.; Haslam, C.; Alexander Haslam, S.; Jetten, J.; Mawson, E.; and Lubman, D. I. 2016. Overcoming alcohol and other drug addiction as a process of social identity transition: The social identity model of recovery (SIMOR). *Addiction Research & Theory*, 24(2): 111–123.

Best, D. W.; and Lubman, D. I. 2012. The recovery paradigm: A model of hope and change for alcohol and drug addiction. *Australian family physician*, 41(8): 593–597.

Betton, V.; Borschmann, R.; Docherty, M.; Coleman, S.; Brown, M.; and Henderson, C. 2015. The role of social media in reducing stigma and discrimination. *The British Journal of Psychiatry*, 206(6): 443–444.

Biegel, D. E.; and Song, L.-Y. 1995. Facilitators and barriers to caregiver support group participation. *J. of Case Management*, 4(4): 164–172.

Blanco, C.; Compton, W. M.; and Volkow, N. D. 2021. Opportunities for research on the treatment of substance use disorders in the context of COVID-19. *JAMA psychiatry*, 78(4): 357–358.

Boughorbel, S.; Jarray, F.; and El-Anbari, M. 2017. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PloS one*, 12(6): e0177678.

Brown, L. D.; Tang, X.; and Hollman, R. L. 2014. The structure of social exchange in self-help support groups: Development of a measure. *American journal of community psychology*, 53(1): 83–95.

Bunting, A. M.; Frank, D.; Arshonsky, J.; Bragg, M. A.; Friedman, S. R.; and Krawczyk, N. 2021. Socially-supportive norms and mutual aid of people who use opioids: An analysis of Reddit during the initial COVID-19 pandemic. *Drug and alcohol dependence*, 222: 108672.

CDC, Centers for Disease Control and Prevention. 2019. Drug overdose deaths. Centers for Disease Control and Prevention website. https://www.cdc.gov/drugoverdose/data/statedeaths.html. Accessed: 2023-01-25.

Chancellor, S.; et al. 2019. Discovering alternative treatments for opioid use recovery using social media. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–15.

Chandrasekharan, E.; Pavalanathan, U.; Srinivasan, A.; Glynn, A.; Eisenstein, J.; and Gilbert, E. 2017. You can't stay here: The efficacy of reddit's 2015 ban examined through hate speech. *Proceedings of the ACM on Human-Computer Interaction*, 1: 1–22.

Choi, M.; Aiello, L. M.; Varga, K. Z.; and Quercia, D. 2020. Ten social dimensions of conversations and relationships. In *Proceedings of The Web Conference 2020*, 1514–1525.

De Choudhury, M.; and De, S. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *Eighth international AAAI conference on weblogs and social media*.

Deri, S.; Rappaz, J.; Aiello, L. M.; and Quercia, D. 2018. Coloring in the links: Capturing social ties as they are perceived. *Proceedings of the ACM on Human-Computer Interaction*, 2: 1–18.

Donovan, D. M.; and Wells, E. A. 2007. 'Tweaking 12-Step': the potential role of 12-Step self-help group involvement in methamphetamine recovery. *Addiction*, 102: 121–129.

D'Agostino, A.; et al. 2017. Social networking online to recover from opioid use disorder: a study of community interactions. *Drug and alcohol dependence*, 181: 5–10.

ElSherief, M.; Saha, K.; Gupta, P.; Mishra, S.; Seybolt, J.; Xie, J.; O'Toole, M.; Burd-Sharps, S.; and De Choudhury, M. 2021. Impacts of school shooter drills on the psychological well-being of American K-12 school communities: a social media study. *Humanities and Social Sciences Communications*, 8(1): 1–14.

Enes, K. B.; Brum, P. P. V.; Cunha, T. O.; Murai, F.; da Silva, A. P. C.; and Pappa, G. L. 2018. Reddit weight loss communities: do they have what it takes for effective health interventions? In *2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 508–513. IEEE.

Eshleman, R.; et al. 2017. Identifying individuals amenable to drug recovery interventions through computational analysis of addiction content in social media. In *2017 IEEE international conference on bioinformatics and biomedicine (BIBM)*, 849–854. IEEE.

Ester, M.; Kriegel, H.-P.; Sander, J.; Xu, X.; et al. 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In *kdd*, volume 96, 226–231.

Etheridge, R. M.; Craddock, S. G.; Hubbard, R. L.; and Rounds-Bryant, J. L. 1999. The relationship of counseling and self-help participation to patient outcomes in DATOS. *Drug and Alcohol Dependence*, 57(2): 99–112.

Eysenbach, G.; and Till, J. E. 2001. Ethical issues in qualitative research on internet communities. *Bmj*, 323(7321): 1103–1105.

Fan, R.; Varol, O.; Varamesh, A.; Barron, A.; van de Leemput, I. A.; Scheffer, M.; and Bollen, J. 2019. The minute-scale dynamics of online emotions reveal the effects of affect labeling. *Nature human behaviour*, 3(1): 92–100.

Fiacco, L.; Pearson, B. L.; and Jordan, R. 2021. Telemedicine works for treating substance use disorder: The STAR clinic experience during COVID-19. *J. of Substance Abuse Treatment*, 125: 108312.

Galanter, M.; White, W. L.; and Hunter, B. 2021. Virtual Twelve Step Meeting Attendance During the COVID-19 Period: A Study of Members of Narcotics Anonymous. *J. of Addiction Medicine*.

Haslam, C.; Dingle, G. A.; Best, D.; Mackenzie, J.; and Beckwith, M. 2016. Social identity mapping: Measuring social identity change in recovery from addiction. In *Addiction, Behavioral Change and Social Identity*, 167–183. Routledge.

Kepner, W.; Meacham, M. C.; and Nobles, A. L. 2022. Types and sources of stigma on opioid use treatment and recovery communities on reddit. *Substance Use & Misuse*, 57(10): 1511–1522.

Krawczyk, N.; Bunting, A. M.; Frank, D.; Arshonsky, J.; Gu, Y.; Friedman, S. R.; and Bragg, M. A. 2021. "How will I get my next week's script?" Reactions of Reddit opioid forum users to changes in treatment access in the early months of the coronavirus pandemic. *Int. J. of Drug Policy*, 92: 103140.

MacLean, D.; et al. 2015. Forum77: An analysis of an online health forum dedicated to addiction recovery. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1511–1526.

McDowall, D.; McCleary, R.; and Bartos, B. J. 2019. *Interrupted time series analysis*. Oxford University Press.

McKeganey, N., James McIntosh. 2000. The recovery from dependent drug use: Addicts' strategies for reducing the risk of relapse. *Drugs: education, prevention and policy*, 7(2): 179–192.

Mead, S.; Hilton, D.; and Curtis, L. 2001. Peer support: a theoretical perspective. *Psychiatric rehabilitation journal*, 25(2): 134.

Mead, S.; and MacNeil, C. 2006. Peer support: What makes it unique. *Int. J. of Psychosocial Rehabilitation*, 10(2): 29–37.

Medvedev, A. N.; Lambiotte, R.; and Delvenne, J.-C. 2018. The anatomy of Reddit: An overview of academic research. *arXiv preprint arXiv:1810.10881*.

Mellis, A. M.; Potenza, M. N.; and Hulsey, J. N. 2021. COVID-19-related treatment service disruptions among people with single-and polysubstance use concerns. *J. of substance abuse treatment*, 121: 108180.

Monti, C.; Aiello, L. M.; De Francisci Morales, G.; and Bonchi, F. 2022. The language of opinion change on social media under the lens of communicative action. *Scientific Reports*, 12(1): 17920.

Moreno, M. A.; Goniu, N.; Moreno, P. S.; and Diekema, D. 2013. Ethics of social media research: common concerns and practical considerations. *Cyberpsychology, behavior, and social networking*, 16(9): 708–713.

Nobles, A. L.; Dreisbach, C. N.; Keim-Malpass, J.; and Barnes, L. E. 2018. " Is This an STD? Please Help!": Online Information Seeking for Sexually Transmitted Diseases on Reddit. In *Twelfth International AAAI Conference on Web and Social Media*.

Ramírez-Cifuentes, D.; Freire, A.; Baeza-Yates, R.; Puntí, J.; Medina-Bravo, P.; Velazquez, D. A.; Gonfaus, J. M.; and Gonzàlez, J. 2020. Detection of suicidal ideation on social media: multimodal, relational, and behavioral analysis. *J. of medical internet research*, 22(7): e17758.

Rapp, R. C.; Xu, J.; Carr, C. A.; Lane, D. T.; Wang, J.; and Carlson, R. 2006. Treatment barriers identified by substance abusers assessed at a centralized intake unit. *J. of substance abuse treatment*, 30(3): 227–235.

Reif, S.; et al. 2014. Peer recovery support for individuals with substance use disorders: Assessing the evidence. *Psychiatric Services*, 65(7): 853–861.

Saha, K.; Kim, S. C.; Reddy, M. D.; Carter, A. J.; Sharma, E.; Haimson, O. L.; and De Choudhury, M. 2019. The language of LGBTQ+ minority stress experiences on social media. *Proceedings of the ACM on human-computer interaction*, 3(CSCW): 1–22.

SAMHSA. Substance Abuse Center for Behavioral Health Statistics and Quality. 2021. Results from the 2020 National Survey on Drug Use and Health: Detailed Tables. https://www.samhsa.gov/data/report/2020-nsduh-detailed-tables. Accessed: 2022-11-01.

Sarason, I. G.; Levine, H. M.; Basham, R. B.; and Sarason, B. R. 1983. Assessing social support: the social support questionnaire. *J. of personality and social psychology*, 44(1): 127.

Schutt, R. K.; and Rogers, E. S. 2009. Empowerment and peer support: Structure and process of self-help in a consumer-run center for individuals with mental illness. *J. of Community Psychology*, 37(6): 697–710.

Sharma, A.; Choudhury, M.; Althoff, T.; and Sharma, A. 2020. Engagement Patterns of Peer-to-Peer Interactions on Mental Health Platforms. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 614–625.

Sowles, S. J.; McLeary, M.; Optican, A.; Cahn, E.; Krauss, M. J.; Fitzsimmons-Craft, E. E.; Wilfley, D. E.; and Cavazos-Rehg, P. A. 2018. A content analysis of an online pro-eating disorder community on Reddit. *Body image*, 24: 137–144.

Stall, R.; and Biernacki, P. 1986. Spontaneous remission from the problematic use of substances: An inductive model derived from a comparative analysis of the alcohol, opiate, tobacco, and food/obesity literature. *Int. J. of the Addictions*, 21(1): 1–23.

Strand, M.; Eng, L. S.; and Gammon, D. 2020. Combining online and offline peer support groups in community mental health care settings: a qualitative study of service users' experiences. *Int. J. of mental health systems*, 14: 1–12.

Strang, J.; Volkow, N. D.; Degenhardt, L.; Hickman, M.; Johnson, K.; Koob, G. F.; Marshall, B. D.; Tyndall, M.; and Walsh, S. L. 2020. Opioid use disorder. *Nature reviews Disease primers*, 6(1): 3.

Tian, Y.; and Chunara, R. 2020. Quasi-experimental designs for assessing response on social media to policy changes. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 14, 671–682.

Tracy, K.; et al. 2016. Benefits of peer support groups in the treatment of addiction. *Substance abuse and rehabilitation*, 7: 143.

Vayshenker, B.; Mulay, A. L.; Gonzales, L.; West, M. L.; Brown, I.; and Yanos, P. T. 2016. Participation in peer support services and outcomes related to recovery. *Psychiatric rehabilitation journal*, 39(3): 274.

Vogel, D. L.; Wade, N. G.; and Hackler, A. H. 2007. Perceived public stigma and the willingness to seek counseling: The mediating roles of self-stigma and attitudes toward counseling. *J. of counseling psychology*, 54(1): 40.

Wright, K. B. 2016. Communication in health-related online social support groups/communities: A review of research on predictors of participation, applications of social support theory, and health outcomes. *Review of Communication Research*, 4: 65–87.

Yang, Z.; et al. 2019. Opioid relapse prediction with gan. In *2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 560–567. IEEE.