Structural Topic Modeling Suggests Importance of Relationships in Addiction Recovery

Sofia Wilson¹ and Dr. Julia Martinez¹#

¹Colgate University
#Advisor

ABSTRACT

Despite the great deal of literature on the problem of heavy drinking and alcohol use, there is relatively little investigation into the lived experience of recovery. The current study addresses this gap through an exploratory thematic analysis of Personal Stories of recovery published in the Alcoholics Anonymous *Grapevine* magazine. Alcoholics Anonymous (AA) is an abstinence-oriented, social support based recovery program. Over 300 Personal Stories were transcribed and analyzed using the Structural Topic Modeling package for R. The model with optimal semantic coherence and exclusivity outlining the five most prevalent topics across all Personal Stories was selected. The five topics were conceptualized and labeled as Relationships and Spirituality, Positive Change, AA and The Fellowship, Addiction and Consequences, and Family. The results of this study suggest social connection to have a strong influence on the recovery process. The topic model also suggests that individuals who find success in sobriety through involvement with AA may truly adopt the principles of the program as their own. Considerations for future directions are discussed.

Alcohol Use Disorder

Alcohol Use Disorder (AUD) is a mentally and physically debilitating condition threatening the safety and well being of individuals, families, and communities across the globe. In 2021, its ruthless grip took hold of 29.5 million lives in the United States (NIAAA, 2023). AUD can manifest regardless of an individual’s race, gender, ethnicity, socio-economic status, and geographic location (NIAAA, 2023). Alcohol-related diseases, injuries, accidents, murders, and suicides steadily remain among the leading causes of death worldwide (NIAAA, 2023). The 25% spike in alcohol-related deaths from 2019 to 2020 warrants consideration of AUD as a public health crisis (NIAAA, 2023). Individuals struggling with AUD experience loss of control over the frequency and quantity at which alcohol is consumed. In the presence of significant social, mental, physical, legal, or other distressing implications, they may desire or unsuccessfully attempt to limit and/or stop their drinking behaviors (American Psychiatric Association, 2017). As the disease perpetuates, changes in brain chemistry are likely to occur in response to prolonged alcohol use, which increases the difficulty of recovery (Witkiewitz et al., 2019).

AUD Treatment

Various empirically supported pharmacotherapies and psychotherapies are available to those suffering from AUD (Witkiewitz et al., 2019). Currently, acamprosate, disulfiram, and naltrexone, are listed as FDA approved medications for AUD (Stokłosa et al., 2023). These drugs can be effective in reducing alcohol consumption by inhibiting the metabolism of alcohol, inhibiting uptake of opioids, and regulating the excitation and inhibition of neural signals (Stokłosa et al., 2023). It is common to treat AUD using a multimodal approach, meaning a combination of medication and psychotherapy. Even with the dramatic increase in prevalence of AUD, conflicting empirical results have thwarted...
identification of a gold-standard psychotherapeutic treatment. The most common psychotherapeutic treatments for individuals with AUD include Cognitive Behavioral Therapy, Dialectical Behavioral Therapy, Motivational Interviewing, and Alcoholics Anonymous or other Twelve Step Facilitation Therapies (Zamboni, 2021; Jhanjee S. 2014; Kelly et al., 2020).

Alcoholics Anonymous

A 2020 Cochrane Review reported the strong likelihood that participation in Alcoholics Anonymous (AA), an abstinence-oriented recovery organization, is the most effective path to recovery (Kelly et al., 2020). For nearly a century, AA has fostered the growth of a global fellowship, which is now estimated to have a membership of over 2 million people (Alcoholics Anonymous, 2023). Member-led meetings address a variety of topics intended to educate, motivate, and support individuals with the desire to stop drinking. The principles of AA, which are outlined in The 12 Steps and 12 Traditions, are adopted by members as a means of structure for their own recovery journey, and guidance for interacting with others in the Fellowship (AA World Services, 2021). The Steps are intended to promote acceptance, making amends, and carrying the message of AA. In more recent years, AA has been used as a model in the creation of alternative Twelve Step Facilitation Therapies. In addition, the material addressed in meetings and in the Alcoholics anonymous big book aim to teach coping strategies, cognitive restructuring, and increase motivation (Morgenstern et al., 1997). AA is unique in the way that members build strong social connections through talking, sharing experiences and struggles, and telling their stories (Robinson et al., 2009).

Recovery in the Literature

Recovering from alcohol addiction is an ongoing process dependent on a combination of complex internal and external factors (Inanlou et al., 2020). However, there are many psychological, social, and environmental factors and processes involved in recovery that are yet to be explored (Witkiewitz et al., 2019, Addison et al., 2020; Best & Hennessy, 2021). An encompassing understanding of the human experience of the addiction recovery process can aid in the development, assessment and implementation of treatment for AUD. Recognizing its importance, existing literature on AUD recovery largely focuses on risk factors and treatment efficacy. Numerous studies on addiction recovery utilize quantitative measures such as frequency and quantity of drinking behaviors, prevalence of psychological symptoms, or qualitative approaches that are lacking in methodological rigor (Glassman et al., 2020). This poses an issue as these measures may limit the depth to which the reality of the addiction recovery experience is explored, or possibly produce biased, ungeneralizable results.

Use of AA and Grapevine Magazine

Alcoholics Anonymous is one of the most frequently pursued, and effective treatments for AUD. The program publishes a monthly magazine, Grapevine, which includes Personal Stories of recovery written and submitted by members of The Fellowship. In these stories, authors detail experiences throughout their own recovery journey, presenting a rich source of qualitative data remaining relatively untapped in terms of research. Grapevine was first utilized for research purposes by Jellinek (1946), who made use of the May 1945 issue to distribute a survey to members of AA. Since then, many criticisms of AA and its members have circulated among scholars. This, along with conflicting empirical results on the program’s efficacy, have created a cloud of controversy surrounding the program and its members. However, many individuals have found the program helpful or essential in their experience of recovery from alcohol addiction.
The Present Study

This study intends to address the lack of research focusing on the lived experience of recovery from alcohol addiction, by performing an exploratory thematic analysis of Personal Stories in the AA *Grapevine* monthly publication. Due to the valued role of storytelling in AA, in addition to the program’s suggested efficacy, these Personal Stories are expected to provide insight into the experienced reality of AUD recovery. This will be performed using The Structural Topic Modeling (STM) package for R, developed by Roberts et al. (2019). This method increases the methodological rigor of this qualitative assessment, by creating a statistically sourced report of the most common topics discussed throughout the narratives. It is hypothesized that AA’s notoriously strong value of connection with both a higher power and other members, will be reflected in the identified topics. This is supported by the program literature including the *Alcoholics anonymous big book* and *The twelve steps and twelve traditions*.

Methods

Participants

A sample of N=300 Personal Stories in AA’s *Grapevine Magazine* from 2012-2023 were transcribed into digital text. These Personal Stories are first-person accounts of experiences with alcohol addiction, recovery and AA involvement. They are written by members of the Fellowship and voluntarily submitted to the monthly publication. There are no restrictions in terms of age, sex, or location restrictions when submitting a story. Although many don’t, authors have the option to keep this information anonymous if their story is published. Personal Stories were submitted from over 40 different states, along with minimal international representation. It is estimated that the authors range in ages from early 20’s to 80’s.

Measures

Three researchers transcribed the 300 stories, while recording a number of associated variables. We demonstrated an extremely high inter-rater reliability, with Cronbach’s alpha of 0.95. The variables of interest included word count, sex of author, publication year, location, years sober, age at decision to get sober, number of quit attempts, presence of family issues, diversity issues and dynamic phrases. Word count is the number of words in the personal story. Sex of the author is defined as either male or female. Years sober represents the length of time that has passed since the author most recently used alcohol. Age at decision to get sober is the age that the author was when they first began their recovery journey. Number of quit attempts is a measure detailing the occurrence of relapse. The presence of family and diversity issues were recorded as present or not present. These included things such as having an alcoholic parent, going through a divorce, or issues due to an individual’s identity such as sexual orientation, disability, or socioeconomic status. Dynamic phrases are phrases that relate to the recovery process. These include things such as higher power, rock bottom, moment of clarity, and one day at a time.

Analysis

*Structural Topic Modeling*

The Structural Topic Modeling (STM) package, developed by Roberts et al. (2019), was used to perform a content analysis on the collection of Personal Stories. I used the package to generate a model based on topic prevalence across the entire corpus of texts. I first processed and prepped the texts which entailed removing any stop words, punctuation, numbers, and any other words that do not appear in between 63-299 stories. Several functions were used to estimate
the appropriate number of topics to be included in the model. A collection of models were generated and evaluated based on exclusivity and semantic coherence of their topics. The best fitting model was selected for interpretation.

**Additional Variables**

In addition to structural topic modeling, the variables we measured along with the Personal Stories were compiled into a dataset. The percentage of stories mentioning family issues and diversity issues were calculated. The average age at which authors decided to become sober, average number of quit attempts mentioned, and average years sober were also calculated. A thematic analysis was performed by looking at the dynamic phrases found across the texts.

**Procedure**

N = 300 Personal Stories in AA’s *Grapevine* Publications from 2012-2023 were transcribed into digital text. Stories were scanned into text, or read out loud and converted to text using Otter.ai. A dataset was created that identified each story along with its associated measures. The STM package for R, as outlined by Roberts et al. (2019), was used to generate a topic model of the Personal Stories. The number of topics to include in the model was determined by assessing held-out likelihood, residuals, and semantic coherence, for models with 3 through 20 topics. The topic number providing the optimal combination of values for these variables was used to generate multiple models. The model offering the best combination of values for semantic coherence and exclusivity, was selected as the model of best fit. The topics within this model were conceptualized and labeled. Lastly, the mean and standard deviation for author sex, age at sober decision, years sober, and word count were calculated. Also calculated was the prevalence of diversity issues and family issues across stories.

**Results**

**Structural Topic Model**

The STM package for R was used to generate a topic model of 300 Personal Stories from AA’s *Grapevine Magazine*. A number of functions were used to determine the best number of topics to include in the model (Figure 1). The model determined as best-fit had five topics, and their respective qualities can be observed in Figure 2. The “FREX” words for each topic represent the words that are highest in frequency and exclusivity for that topic, and the “highest prob” words are the words that show up most often in that topic but are not exclusive to that topic (Figure 3). Topics were interpreted and labeled as follows: Topic 1: Family, Topic 2: AA and The Fellowship, Topic 3: Positive Change, Topic 4: Relationships and Spirituality, Topic 5: Addiction and Consequences. Word cloud representations of each topic can be seen in Figure 4.
Figure 1. Determining the Best Number of Topics. This figure shows the values for held-out likelihood, semantic coherence, residuals, and lower bound based on different numbers of topics in the model. The searchk() function was used to produce these graphs.

Figure 2. Quality of Topics for Selected Model. This figure shows the quality of each topic in the selected model of best fit. Topic quality is based on semantic coherence and exclusivity of the words within the topic. The topicQuality() function was used to produce this graph.
Figure 3. Most Frequent and Highest Probability Words for Topics in Best-fit Model. This figure shows the 13 highest probability and FREX words for each topic. These are the words used to conceptualize and label each topic. Highest probability words are the words that are most likely to appear when that topic is being discussed. These words are not exclusive to the topic. The FREX words are the words with the highest frequency in that topic, and are exclusive to that topic. The FREX words are most telling about the content of the topic, however highest probability words can be used for additional conceptualization. The labelTopics() function was used to obtain this figure.

Figure 4. Word Cloud Representations of Topics 1-5. Word size in the word cloud is determined by its frequency of use. Large words indicate high prevalence within that topic. The cloud() function was used to produce this graphic.
Relationships and Spirituality is the topic with the highest prevalence, accounting for 23% of all text. Family is the topic with the lowest prevalence, accounting for 16% of all text. The remaining three topics fall in between. When looking at all text together, topics are discussed in fairly equal amounts (Figure 5), but this does not mean all individual stories discuss each topic evenly (Figure 6). Correlations between topics indicate the likelihood that they are discussed in the same individual story. Addiction and Consequences, Positive Change, and Family were positively correlated with each other. An additional positive correlation was found between Relationships and Spirituality and AA and The Fellowship (Figure 7).

**Figure 5.** Topic Proportions. This figure shows the expected proportion of text dedicated to each topic. The topic most prevalent among the 300 Personal Stories is Relationships & Spirituality followed by Positive Change, AA & The Fellowship, Addiction and Consequences, and lastly Family. The plot () function was used to obtain this graphic.

**Figure 6.** Distribution of Topics. This figure shows the distribution of topic proportions across all individual stories. On average, Topic 1 makes up 16% of each individual story, Topic 2 makes up 18% of each individual story, Topic 3 makes up 22% of each individual story, Topic 4 makes up 23% of each individual story, and Topic 5 makes up 16% of each individual story. These values
are the average prevalence across all stories. In each individual story, the topics are not as balanced. The plot(type="hist") function was used to obtain this graphic.

---

**Figure 7.** Topic Correlations. This figure demonstrates the correlational relationships in terms of prevalence between all five topics. Positive correlations between topics indicate a high likelihood that they are discussed in the same individual story. Topics 1, 3 and 5 are positively correlated with each other, while topics 2 and 4 are positively correlated with each other. The topicCorr() function was used to obtain this graphic.

---

**Dataset Analysis**

The data analysis revealed the mean word count for the Personal Stories is 1030.37 words (SD= 383). About a third of the stories mention family issues (Figure 8), while only 11.33% of the Personal Stories mention diversity issues (Figure 8). Almost two thirds, 66.33%, of the Personal Stories were written by women (Figure 8). The average age at which the authors decided to get sober was 33.71 years old (SD= 12). The average length of sobriety is 19.46 years (SD= 13.47), and the average number of quit attempts mentioned is 1.28 (SD= 1.29).

---

**Table 1. Mean Values for Additional Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Count</td>
<td>1030.37</td>
<td>383.66</td>
</tr>
<tr>
<td>Age at Sober Decision</td>
<td>33.71</td>
<td>12.20</td>
</tr>
<tr>
<td>Years Sober</td>
<td>19.46</td>
<td>13.47</td>
</tr>
<tr>
<td>Quit Attempts Mentioned</td>
<td>1.28</td>
<td>1.29</td>
</tr>
</tbody>
</table>

This table shows the mean and standard deviation for Wordcount, Age at Sober Decision, Years Sober and Quit Attempts Mentioned across all N=300 Personal Stories.
Figure 8. This figure shows the percentage of Personal Stories that mention family issues (32.33%), and diversity issues (11.33%). Also in this figure is the percentage of stories that were written by men and women (36.33% male, 63.37% female).

Discussion

Sharing stories and experiences in addiction and recovery is central to the AA program. Meetings are structured to encourage individuals to share their own stories and learn from others’. The Personal Stories in Grapevine are a rich source of qualitative data, detailing individuals’ lived experience of recovery from alcohol addiction. Topics highly prevalent across these stories may be indicative of changes that occur in the recovery process. A deeper understanding of the lived experience of recovery can inform treatment efforts, and relapse prevention.

The Topic Model

The Structural Topic Modeling package for R (Roberts et al., 2019), was used to perform a content analysis of 300 Personal Stories of Recovery from Alcohol Addiction. The stories were written by members of Alcoholics Anonymous and published in the Grapevine magazine. The five topics that make up the model determined to be of best fit were Family, AA and The Fellowship, Positive Change, Relationships and Spirituality, Addiction and Consequences. When looking at all texts together, the difference in prevalences of the five topics is not drastic. This suggests each of these topics to have a relatively similar level of importance when looking across all stories. However, this does not mean that authors discussed each topic in equal amounts, as most stories focus on one or two of these topics.

Relationships and Spirituality

Relationships and Spirituality was the topic with the highest prevalence, meaning it was discussed the most across all Personal Stories. This supports the hypothesis posed in this study, which was that AA’s notoriously strong value of connection with both a higher power and other members, will be reflected in the identified topics. This means that for individuals in AA, these two things may have a strong influence on their recovery process. Considering the strong spiritual component of AA, this is understandable. In the Alcoholics anonymous big book and The twelve steps and twelve traditions, individuals are encouraged to connect with, and be guided by their Higher Power. Some of the most common words associated with this topic include: help, God, sponsor, book, love, read, and step; which suggest a positive impact of relationships and spirituality on the recovery process. Two of the most common dynamic phrases identified in the stories were “Higher Power” and “God”. It makes sense that spirituality would play an important role in the recovery of those in AA, as it is largely involved in the program’s teachings and principles. The relationship aspect of this topic may likely be describing the new friendships and mentorships formed through the Fellowship. Many of the Personal Stories described a newfound sense of belonging and community once becoming involved in AA. AA encourages members to reach out to each other for support. At their first meeting, an individual will receive a book of names and phone numbers of members they call in times of crisis. The contacts in the book include their information voluntarily. Many of The Twelve Steps address aspects of life that have to do with healing relationships.
and making amends to those who had been harmed by an individual’s drinking. These common practices demonstrate the emphasis AA places on social support/connection during recovery. It would make sense that the relationships formed as a result would play a significant role in an individual’s recovery.

Positive Change
Positive Change is the topic with the second highest prevalence across all of the Personal Stories. Some of the words associated with this topic include: change, remember, good, new, look, back, can. This topic makes sense to be of significant prevalence in the Personal Stories, as these stories describe the journey from addiction to sobriety. These words are likely used often in the text in descriptions of sobriety compared to some earlier time in an individual’s life when their addiction was more active. Two dynamic phrases commonly used when describing the experience of addiction include: “desperation” and “rock bottom,” while “saved my life,” and “gratitude” were often used to describe life once sober. When reading the Personal Stories, it was apparent that quality and enjoyment of life improved dramatically once recovery came into the picture. This, along with the use of dynamic phrases and the words associated with this topic, support the idea that positive change is a common experience in the recovery journey of the authors.

AA and The Fellowship
AA and The Fellowship was the third most prevalent topic across all of the Personal Stories. Words frequently appearing when discussing this topic include: meet, member, group, experience, share, power, recovery, year, time. Similar to Relationships and Spirituality, these words indicate the sense of community in AA, reported by many of the authors. AA and The Fellowship were positively correlated with Relationships and Spirituality. Meaning if one of the topics is present in an individual story, the likelihood of the other being discussed as well is higher. This topic likely appears when the authors are describing their involvement in the program. For example, how often they go to meetings, how long they have been a member, or specific experiences such as sharing at a meeting or listening to other members. This topic is not surprising considering all of the authors have/had some degree of involvement in AA, and Grapevine is the program’s publication.

Addiction and Consequences
Addiction and consequences was the fourth most prevalent topic across all Personal Stories. This topic was interpreted from words including: drink, drunk, couldn’t, stop, problem, knew, alcohol, and job. It is likely that this topic appears in the text when authors discuss experiences in active addiction. It was noticed that many Personal Stories followed a similar structure. First they described how and when they got addicted to drinking, how the disease progressed and began to affect their daily life, how they ended up in AA, and lastly, how their life has changed as a result of sobriety. Common phrases used to describe life during active addiction include: “powerless” “suicide” “misery” and “hopeless.” These phrases support the idea that addiction negatively impacted the lives of these authors. There appeared to be a trend where authors realized the extent of the consequences of their addiction, which evoked motivation to enter recovery. A lot of the time these consequences included things like losing a job, divorce, or physical and mental health struggles. When this happened, it was often referred to as rock bottom.

Family
Family was the fifth most prevalent topic across all Personal Stories. This means it was the least discussed out of the five topics, however, still accounts for a significant amount of the text. It does not come as a surprise that this is one of the topics identified through STM, as addiction is often referred to as a family disease. This means it affects many more people than just the individual who is addicted. Some of the words associated with this topic include: wife, family, mother, house, children, husband, father, and without. Many of the authors mentioned family members who also struggled with addiction. The negative effects of drinking on family members, and losing family members due to addiction were also frequently mentioned. The presence of family issues was one of the variables measured when
transcribing the Personal Stories. However, only a third of the stories indicated family issues. This means it is unlikely that all stories that discuss family, do so in a negative way. This makes sense as many stories described how certain family members were crucial in terms of realizing drinking had become a problem, and/or getting help to begin recovery. The topic of Family was also positively correlated with Addiction and Consequences, and Positive Change. This supports the idea that family can be involved in an individual’s recovery in both positive and negative ways.

**Important Considerations**

It is important to keep in mind that structural topic modeling is an exploratory approach to analyze the content of large amounts of text. While performing the analysis, it became apparent that the outcome is heavily reliant on the interpretation and evaluation of the researcher. STM produces a number of models, but it is up to the researcher to make decisions along the way and ultimately decide which model is the best fit. Researcher decisions included: identifying the best number of topics to be included in the models, setting thresholds for word exclusion, and assessing model and topic quality. There is no right or wrong outcome with this approach to thematic analysis, but it can change depending on the choices made by the researcher along the way. STM may get rid of some interpretation bias which is a common concern when doing qualitative research, but it does not completely eliminate it.

**Strengths, Limitations and Future Directions**

**Limitations**

This study should be taken in the context of its limitations. For one, there is still concern for bias from the researcher when evaluating topics and models. In addition, the Personal Stories in AA’s *Grapevine* Magazine are not required to include any specific information. The stories are written by individuals in recovery, where they can include, or not include whatever information they choose. This means certain experiences in the recovery process could be left out. This is of importance especially in the case of the additional variables measured during transcription. Age at sober decision, quit attempts, and years sober were not always explicitly mentioned by the author. Many of the times we had to do some small calculations and/or estimations based on other dates mentioned. So, the results of these calculations may not be completely representative of the true values for these measures. However, it is possible that the absence of content requirement makes the results of the topic modeling even more telling. It is likely that the authors chose to include the aspects of their recovery journey in their Personal Stories that were the most influential. In addition, there were multiple researchers involved in identifying dynamic phrases, and collecting the additional variables. This does raise concern for inter-rater reliability, as there was no operational definition of dynamic phrases established before data collection. However, the inter-rater reliability was extremely high, suggesting consistency between all three researchers when measuring additional variables and determining dynamic phrases ($\alpha = 0.95$).

**Strengths**

This study also offers a number of strengths. A thematic analysis of this size, using personal narratives of alcohol addiction recovery has yet to be conducted. STM is a relatively new method, and has not yet been used to analyze such rich qualitative data on this fairly under researched topic. STM does take away some of the bias traditionally associated with qualitative research. Simply reading and looking for themes in such a large amount of text (over 300,000 total words) introduces the threat of confirmation bias and researcher fatigue. STM takes a lot of this possibility away as it uses probabilities and frequencies of words and phrases to determine what is most prevalent. In addition, many empirical studies on addiction and addiction recovery restrict their participant groups to people with DSM diagnoses of alcohol addiction, and no comorbidities. This may not be generalizable to the greater public, as alcohol addiction is a disease with an extremely high rate of comorbidity. AA however is an extremely accessible program that only requires the desire to stop drinking for membership. Many authors of the Personal Stories included mention of other mental health struggles.
**Future Directions**

This study offers ample opportunity for future direction. The next step in this project is to look for correlations between the generated topics and the additional variables that were measured (age, sex, years sober, number of quit attempts). This can be performed using the topic Corr() function in STM. In addition, other models should be explored. Some ways to generate new models include using different numbers of topics, different upper and lower thresholds and specifying certain words to be excluded. There is also opportunity to transcribe more Personal Stories into the corpus of text. Another future direction would be to perform an STM analysis on interviews. In this way, the richness of the personal narrative is preserved while introducing the possibility for the interviewer to collect additional variables including age at decision to get sober, years sober, presence of family and/or diversity issues, and number of quit attempts.

**Conclusion**

This was an exploratory study aiming to identify themes in the lived experience of recovery from alcohol addiction. A novel method to approach this topic was used and revealed five topics prevalent across the Personal Stories. It is important to recognize that addiction recovery is a variable process, experienced differently by every individual. However, the results of this study signify a shared strong influence of social connection in the recovery process. AA is heavily reliant on relationships, connection and social support. It may be the case that members who find success in sobriety as a result of AA involvement, truly immerse themselves into the program and adopt the principles as their own. Regardless, the criticisms of AA stemming from the fact that it is not a professionally, or clinically run program shouldn’t thwart it from being used as a model of recovery in research. This is not an argument that it be deemed the gold standard approach as this program is not without its faults and does not work for everyone (Glassman et al., 2023). However, many individuals have reported the program to be helpful, or essential, to their recovery journey. This alone suggests reason for further investigation of the potential mechanisms of change experienced with involvement in AA.

**References**


Alcoholics Anonymous World Services, Inc. (2021). *Twelve steps and twelve traditions*.


