Making Sense of Qualitative Data Using Interactive Network Visualization: A Case Study of Narratives Concerning Stigma Relating to Substance Use

Annie T. Chen* University of Washington

Lexie C. Wang† University of Washington

Bogeum Choi‡ University of North Carolina Chapel Hill

Adam Babitts§ University of Washington

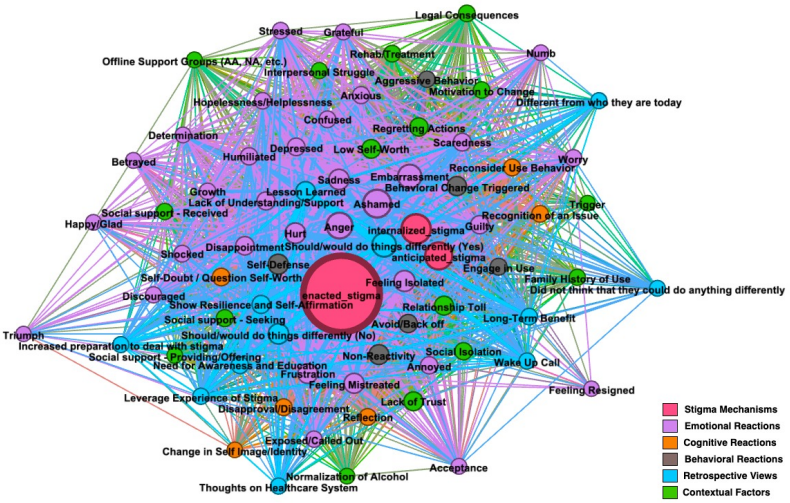

Figure 1: Thematic Network in a Survey of Substance Use-Related Stigma Experiences

Abstract

Qualitative data analysis is a process involving the extraction of meaning from a sample of artifacts. While this process can yield rich insights, there can also be substantial complexity, and a need for greater analytical capabilities than afforded by a given qualitative data analysis tool. In this poster, we present a novel pipeline for qualitative data analysis that facilitates investigation of complex psycho-social phenomena. We illustrate our use of this pipeline collected to analyze survey data, to elucidate the mechanisms at work in stigma relating to substance use.

Index Terms: Network visualization, qualitative data analysis, thematic analysis, stigma, substance use.

1 Introduction

Many approaches to qualitative data analysis (QDA) involve researchers examining text and assigning descriptive codes to it. QDA software can facilitate interpretive practices including but not limited to: exploration of multiple meanings and gaps in the data, viewing the data through particular conceptual lenses, challenging researcher assumptions, and reflecting on the social construction of evidence [8]. However, in the extant literature there is a lack of recognition and discussion of the role that software can play in interpretive practice, and challenges that may arise [8, 19].

QDA software tools such as Atlas.ti, Dedoose, and MAXQDA offer analytics features such as code application matrices, co-occurrence matrices, and more. However, there can be limitations; for example, a code application or code co-occurrence matrix could work well for interviews involving smaller samples, but exhibit slow performance or poor interpretability for large datasets.

To this end, we might consider other approaches such as distant reading. A term first coined by Moretti [14], distant reading involves shifting from close reading of text, to deriving abstract views of it by visualizing features within and potentially across multiple texts, to glean meaning [11]. These abstractions may involve heatmaps, tag clouds, maps, graphs, timelines, and others. Since the introduction of distant reading, there has been some debate, with close and distant reading often being depicted as existing in opposition [9]; others argue that both are important and should be employed together [3, 6], with distant reading giving rise to insights that can be investigated further through close reading [6].

We support the position that both have their respective contributions to analysis, and present a case study involving qualitative data analysis of survey data concerning an important health issue – stigma relating to substance use – to illustrate how QDA tools might be combined with visualization tools to facilitate interactive exploration, transparency, and rigor in qualitative analysis.

2 Visual Analysis Pipeline

The analytic process includes the following parts: data collection, pre-processing, derivation of variables from text data through the process of qualitative coding, processing, and subsequent interactive discovery activities (Figure 2). After performing qualitative coding, the codes are exported from the QDA software (in this case, Dedoose), processed to facilitate subsequent analysis through data transformations such as aggregation, and then used in interactive exploration and inquiry, including statistical and visual analyses. In this poster, we explore the use of the pipeline to facilitate network visualization in Gephi v.0.10.1 [1], which are used in concert with other visualizations to facilitate complementary views of the data.

*e-mail: atchen@uw.edu

†e-mail: lexwang@uw.edu

‡e-mail: choiboge@unc.edu

§e-mail: ababitts@uw.edu

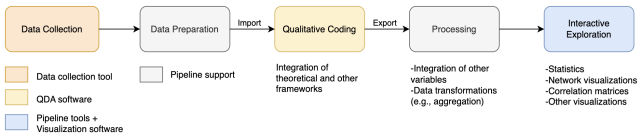

Figure 2: QDA thematic analytic pipeline.

3 Data Description, Representation, and Analysis

Stigma is a social phenomenon that can involve labeling, stereotyping, separation, status loss, and discrimination due to a stigmatized attribute [13]. People may experience stigma due to various health conditions. In this project, we explore stigma relating to substance use because of the far-reaching consequences that it has on different facets of life, including physical and mental health, housing stability, and willingness to seek treatment [5, 12]. We conducted a survey, recruiting participants via Prolific [15]. We asked participants to share narrative responses of their experiences (often in the form of incidents) of substance use stigma. Four coders performed qualitative data analysis of the responses (N=510).

We developed a code hierarchy incorporating multiple theoretical frameworks. For example, this research employs the Stigma Framework [7], which posits three stigma mechanisms: internalized (self-stigma), anticipated (anticipation of stigma from others), and enacted stigma (discrimination, stereotyping etc. based on a stigmatized attribute) [17]. As the extant literature has observed that there is a lack of clarity concerning how stigma mechanisms impact recovery [5], we defined various code families, including emotional, cognitive and behavioral reactions, to explore and articulate how the stigma mechanisms operate. We developed a nomenclature to ensure that the codes appeared in a logical structure in the QDA software; this data representation is retained in subsequent visualization steps to facilitate research inquiry.

3.1 Using Network Visualization to Articulate Psycho-Social Phenomena

One challenge in studying social phenomena is the diverse factors to be considered in a non-linear and multi-factorial fashion. Network analysis can facilitate exploration of complex relationships between codes, as well as enhanced transparency in the analytic process [16]. We present sample network visualizations that we employ in concert with other visualizations in analysis.

We employ graph visualizations in which thematic codes are represented as nodes. Their co-occurrence in the responses are represented as edges, or connections between the nodes. Nodes are sized by their frequency, with larger nodes representing more common codes, and colorized by the code families. In Figure 1, the most common theme is the stigma mechanism, enacted stigma (pink).

The network layout influences how the codes appear spatially, and hence, how one might subsequently interpret them. We employ ForceAtlas2 [10], a common layout algorithm that leverages energy attraction and repulsion in the network to approximate relation strength. At a high level, the concepts that tend to co-occur in the same survey responses (or incidents described by respondents) will generally appear closer in the visualization.

Stigma can be associated with varied emotions (Figure 1, purple) and behavioral reactions (grey), including self-defense, engaging in substance use, and avoidance. Participants also mentioned cognitive reactions (orange), including reflection, recognition of issues, and questioning their own self-worth. The visualization illustrates a variety of different reactions to a stigma experience, which could in turn be used to develop strategies to facilitate adaptive coping. Filtering the network visualization can reduce complexity and potentially facilitate interpretability, though it is important to consider whether the tradeoff is appropriate for the research question [4].

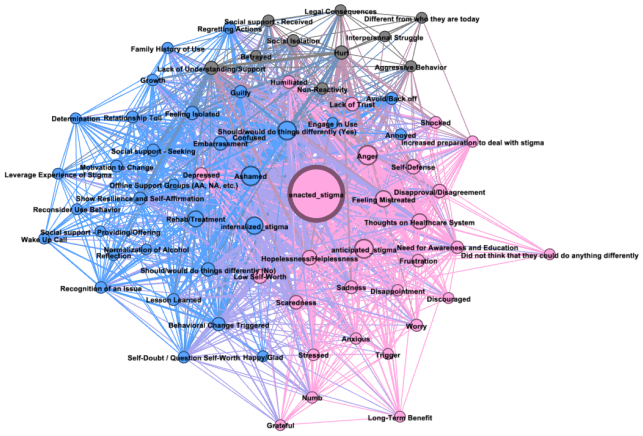

Figure 3: Common reactions in healthcare contexts: Thematic groupings identified using the Louvain method [2].

3.2 Interactive Exploration of Particular Contexts

The pipeline facilitates interactive exploration of research questions through different lenses and perspectives. In this case, we sought to examine how stigma mechanisms manifest depending on the setting (e.g., home, work, healthcare, leisure) and substance (i.e., alcohol, cannabis, opioids). We subset and filter the data by qualitative codes (e.g., setting) and structured data (e.g., race, comorbidities), re-rendering visualizations and re-computing statistics as needed.

Stigma relating to substance use is a recognized concern in healthcare [18]. To explore the shared experiences of persons who engage in substance use, we employed subcommunity detection using the Louvain method [2]. Different levels of aggregation enable exploration of alternate views: how the factors in a single incident (incident-level) and across incidents (person-level) may influence experience. In Figure 3, we employ incident-level aggregation to more clearly depict the factors in one setting, healthcare, than if we allowed for blurring across life spheres. This is an example of one visualization that we employ while performing QDA along with other visualizations and the underlying responses, to understand the complex relationships between the variables.

One thematic grouping (pink) included enacted stigma, which could be accompanied by anger, feeling mistreated, and hopelessness. This grouping also included perspectives on healthcare, a need for awareness, and anticipated stigma, representing survey responses referencing difficult healthcare encounters leading to avoidance of future medical treatment. Another thematic grouping (blue) included internalized stigma, which was associated shame, embarrassment, and confusion. Working with the visualization interactively and considering the connections, we also observed that these experiences occasionally gave rise to transformative processes such as recognition of issues and motivation to change. Other reactions, such as feeling hurt due to a lack of understanding and support from others, were depicted in a third grouping (grey). These reactions, which might occur both between and within individuals (at different times), could inform intervention development.

3.3 Conclusion

We illustrate how an analytic pipeline integrating QDA and visualization tools can facilitate exploration of multidimensional relationships between codes. Though conventional QDA tools facilitate rich analyses, leveraging additional tools can further augment the analytic process. Our approach opts to retain complexity as opposed to reducing it, but we recognize the inherent challenges. Multiple visualizations and views also facilitate triangulation to ensure research validity.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Institute on Drug Abuse of the National Institutes of Health under Award Number R21DA056684.

REFERENCES

[1] M. Bastian, S. Heymann, and M. Jacomy. Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, 3(1):361–362, Mar. 2009. doi: 10.1609/icwsm.v3i1.13937 1

[2] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, Oct. 2008. doi: 10.1088/1742-5468/2008/10/P10008 2

[3] R. Bonfiglioli and F. Nanni. From Close to Distant and Back: How to Read with the Help of Machines. In F. Gadducci and M. Tavosanis, eds., *History and Philosophy of Computing*, vol. 487, pp. 87–100. Springer International Publishing, Cham, 2016. Series Title: IFIP Advances in Information and Communication Technology. doi: 10.1007/978-3-319-47286-7_6 1

[4] A. T. Chen and C. L. Cole. Reflexivity in issues of scale and representation in a digital humanities project. In *6th Workshop on Visualization for the Digital Humanities*, p. 7, Oct. 2021. doi: 10.48550/arXiv.2109.14184 2

[5] K. A. Crapanzano, R. Hammarlund, B. Ahmad, N. Hunsinger, and R. Kullar. The association between perceived stigma and substance use disorder treatment outcomes: a review. *Substance Abuse and Rehabilitation*, 10:1–12, Dec. 2018. doi: 10.2147/SAR.S183252 2

[6] Drouin. Close- and distant-reading modernism: Network analysis, text mining, and teaching the Little Review. *The Journal of Modern Periodical Studies*, 5(1):110, 2014. doi: 10.5325/jmodeperistud.5.1.0110 1

[7] V. A. Earnshaw and S. R. Chaudoir. From Conceptualizing to Measuring HIV Stigma: A Review of HIV Stigma Mechanism Measures. *AIDS and behavior*, 13(6):1160–1177, Dec. 2009. doi: 10.1007/s10461-009-9593-3 2

[8] L. S. Gilbert, K. Jackson, and S. Di Gregorio. Tools for Analyzing Qualitative Data: The History and Relevance of Qualitative Data Analysis Software. In J. M. Spector, M. D. Merrill, J. Elen, and M. J. Bishop, eds., *Handbook of Research on Educational Communications and Technology*, pp. 221–236. Springer New York, New York, NY, 2014. doi: 10.1007/978-1-4614-3185-5_18 1

[9] N. Glaubitz. Zooming in, zooming out: The debate on close and distant reading and the case for critical digital humanities. In A.-J. Zwierlein, J. Petzold, K. Boehm, and M. Decker, eds., *Anglistentag 2017*, pp. 21–28. WVT, Trier, 2018. 1

[10] M. Jacomy, T. Venturini, S. Heymann, and M. Bastian. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PLOS ONE*, 9(6):e98679, June 2014. doi: 10.1371/journal.pone.0098679 2

[11] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann. Visual Text Analysis in Digital Humanities. *Computer Graphics Forum*, 36(6):226–250, 2017. doi: 10.1111/cgf.12873 1

[12] M. Kulesza, S. Ramsey, R. Brown, and M. Larimer. Stigma among Individuals with Substance Use Disorders: Does it Predict Substance Use, and Does it Diminish with Treatment? *Journal of addictive behaviors, therapy & rehabilitation*, 3(1):1000115, Jan. 2014. doi: 10.4172/2324-9005.1000115 2

[13] B. G. Link and J. C. Phelan. Conceptualizing Stigma. *Annual Review of Sociology*, 27:363–385, 2001. Publisher: Annual Reviews. 2

[14] F. Moretti. *Distant Reading*. Verso, London, 1st edition ed., June 2013. 1

[15] E. Peer, D. Rothschild, A. Gordon, Z. Evernden, and E. Damer. Data quality of platforms and panels for online behavioral research. *Behavior Research Methods*, 54(4):1643–1662, Sept. 2021. doi: 10.3758/s13428-021-01694-3 2

[16] J. J. Pokorny, A. Norman, A. P. Zanesco, S. Bauer-Wu, B. K. Sahdra, and C. D. Saron. Network analysis for the visualization and analysis of qualitative data. *Psychological Methods*, 23(1):169–183, Mar. 2018. doi: 10.1037/met0000129 2

[17] L. R. Smith, V. A. Earnshaw, M. M. Copenhaver, and C. O. Cunningham. Substance Use Stigma: Reliability and validity of a theory-based scale for substance-using populations. *Drug and alcohol dependence*, 162:34–43, May 2016. doi: 10.1016/j.drugalcdep.2016.02.019 2

[18] L. C. Van Boekel, E. P. Brouwers, J. Van Weeghel, and H. F. Garretsen. Stigma among health professionals towards patients with substance use disorders and its consequences for healthcare delivery: Systematic review. *Drug and Alcohol Dependence*, 131(1-2):23–35, July 2013. doi: 10.1016/j.drugalcdep.2013.02.018 2

[19] P. Zhao, P. Li, K. Ross, and B. Dennis. Methodological Tool or Methodology? Beyond Instrumentality and Efficiency With Qualitative Data Analysis Software. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 17(2), Apr. 2016. Number: 2. doi: 10.17169/fqs-17.2.2597 1