Qualitative Psychology

Manuscript version of

Using Text-Analysis Computer Software and Thematic Analysis on the Same Qualitative Data: A Case Example

Ruth L. Firmin, Kelsey A. Bonfils, Lauren Luther, Kyle S. Minor, Michelle P. Salyers

_____

Abstract

The acceptance and application of qualitative methods has been steadily increasing, and recent advances in computer analytic software programs have produced a rapidly-evolving landscape of new methods and analytic tools. However, discussions regarding the use of these new computer-based methods alongside traditional qualitative methods remain sparse. The aim of this article is to present an example of using quantitative text analysis software, the Linguistic Inquiry and Word Count program (LIWC), alongside a traditional qualitative method, thematic analysis. Data included 46 transcribed life-narratives shared by individuals with schizophrenia. We present findings from both analyses and offer an example of a method that combines these two approaches. Results and examples provided are discussed in light of the potential to strengthen analyses by using these methods collaboratively.

Keywords

Using Text-Analysis Computer Software and Thematic Analysis on the Same Qualitative Data:

A Case Example

The use of qualitative methods has been increasing over the past several decades and there are more qualitative data techniques avaliable to researchers than ever before (Atchison, 1996; Creswell, 2012). Moreover, the tools for qualitative analyses have also been expanding rapidly (La Pelle, 2004; Meyer & Avery, 2009). Today, text-analysis software programs offer the ability to quantify qualitative data more efficiently. Electronic assistance to qualitative coding has grown from using functions available in programs like Microsoft Word and Excel (e.g., word count), to using these and other programs to organize and assist with qualitative coding and memos (La Pelle, 2004). Today, qualitative software programs abound that offer researchers assistance with organizing and tracking qualitative coding, the ability to rapidly pull text labeled with a given code across transcripts, or the ability to store and analyze visual or other forms of data, among other capabilities (Creswell, 2012). Moreover, electronic coding tools are being used in studies that employ inductive coding methods, including grounded theory (Bringer, Johnston, & Brackendidge, 2006; Firmin, Luther, Lysaker, Salyers, 2015).

One important recent advancement has been the ability for text-analysis software programs to produce more detailed textual analysis, providing researchers with information about the types of words used, the types of categories into which words fall, and the types of topics most represented in given data. Several dictionary-based content analysis programs exist that offer computer-assisted qualitative data analysis. In 2002, Lowe reviewed qualitative data analysis software and noted key differences between software designed to assist the researcher with organizing and annotating (e.g., Atlas-ti, NUDIST) compared to software designed to assist with content analysis. At this time, Lowe noted 15 programs that used dictionary-based content

analysis. One program that has been widely used in the social sciences and other fields is the

Linguistic Inquiry and Word Count program (LIWC), which calculates the percent of one's

speech that falls within various domains that are meaningful and relevant to psychological

processes, such as thinking styles, emotion, and social relationships (Pennebaker, Booth, &

Francis, 2007; Buck, Minor, & Lysaker, 2015; Buck, Minor, & Lysaker, in press; Minor et al.,

2015). The LIWC has been validated and found to produce reliable groupings of words that

relate to language and other relevant characteristics such as participant variables in expected

ways (Pennenbaker et al., 2007; Minor et al., 2015).

These recent advances offer researchers the ability to conduct more nuanced analyses

with increased sophistication. Indeed, these programs offer several notable advantages. First,

many mundane aspects of qualitative coding, such as word-counting, that may have been

impractical using traditional methods, can now be conducted quickly using these new software

programs (John & Johnson, 2004). Moreover, these same authors note that researchers are now

able to analyze larger quantities of qualitative data than before, yet also warn that pitfalls of these

methods could include losing focus on depth and meaning that qualitative data can offer.

With the growth of qualitative analytic tools and approaches, researchers have advocated

the use of multiple analytic approaches, so as to capitalize on the strengths offered by each tool

(Johnson & Onwuegbuzie, 2004; Leech & Onwuegbuzie, 2007). Indeed, triangulation, or using

multiple data sources to examine a phenomenon or construct, has since been discussed as a key

method of enhancing validity and reliability of qualitative analyses (Cresswell, 2014). One

approach to triangulation is through employing mixed-methods designs (Bryman, 2006). Mixed-

methods approaches integrate qualitative and quantitative data and this approach is becoming

increasingly popular for its potential to draw from the strengths of both qualitative and

quantitative approaches to data analysis and thereby potentially minimize the weaknesses accompanying both approaches (Cresswell, 2014). Less work has discussed the role of more advanced qualitative data analysis tools regarding mixed-methods designs. Indeed, although newer text-analysis tools have the capacity to conveniently and reliably analyze qualitative text in a quantitative manor, relatively little work exists discussing how these new tools may fit into the traditional qualitative methods landscape. Furthermore, little guidance currently exists for researchers interested in taking advantage of the benefits of these sophisticated tools but who also wish to draw from the benefits of traditional thematic analyses.

This paper aims to address this gap in the emerging literature by offering a case example of using computer text-analysis software and thematic qualitative analyses on the same data. Findings using both approaches are presented and compared. We then present an approach to using the methods in conjunction with each other. We used thematic analysis to illustrate a widely-used, traditional qualitative method that relies on an inductive method and that is relatively work-intensive (Cresswell, 2013). To illustrate a text-analysis program, we selected the LIWC, given its increasing popularity, relative affordability and rapid analyses, capacity to provide more nuanced categories of words that reflect psychological processes, and its prior use with similar populations (Buck & Penn, 2015; Cohen, et al., 2007; Pennebaker, Booth, & Francis, 2007).

Method

Setting and Participants

Data for both thematic and LIWC analyses were obtained for 46 participants enrolled in a larger randomized controlled trial of illness management and recovery (IMR), a curriculum-based illness self-management program for mental illness (Gingerich & Mueser, 2005). This

group was recruited to participate in a sub-study for which they completed an in-depth narrative interview prior to receiving the IMR intervention. Participants were eligible for the study if they spoke fluent English and were receiving mental health services at either the VA Medical Center or a local community mental health center, were older than 18 years of age, had a diagnosis of a schizophrenia-spectrum disorder (as confirmed by the Structured Clinical Interview for the DSM-IV; First, Spitzer, Gibbon, & Williams, 1996), were able to pass a cognitive screener (Callahan, Unverzagt, Hui, Perkins, & Hendrie, 2002), and had no physical health condition that would prevent participation in an 18 month longitudinal study. The majority of participants were men ($n = 35$, 76.1%) and Black ($n = 29$, 63.0%). Most participants had completed high school or had gone on to additional education ($n = 30$, 65.2%). The mean age of participants was 48.7 years (SD = 8.7).

Procedures

Interviews were conducted by trained research assistants and typically lasted less than one hour. The interview protocol, the Indiana Psychiatric Illness Interview (IPII; Lysaker et al., 2005), was a semi-structured life narrative interview that has been widely used among persons with serious mental illnesses, including schizophrenia-spectrum disorders, to elicit perceptions regarding one's self and illness (Lysaker, Clements, Plascak-Hallberg, Knipscheer, & Wright, 2007). The interview consisted of rapport building and open-ended questions that invite individuals to first share their life story and then discuss their perceptions of their illness and its impact on their life. Question stems are consistent across participants, and interviewers were trained to only probe occasionally, intentionally letting participants share their life story and perceptions about their illness in ways that were meaningful to them. All participants provided informed consent and were compensated $20. Procedures were approved by the Institutional

Review Boards at the VA and the university. For a full description of the recruitment procedures, see Firmin, Luther, Lysaker, & Salyers (in press). All interviews were audio recorded, transcribed verbatim, and de-identified. The same interview transcripts were used for both the thematic analysis, and the LIWC analysis, as described below.

Analyses

The qualitative methods were conducted using a thematic analysis approach. Our method was similar to a modified grounded theory approach in that we employed an inductive, iterative approach to generating themes and codes that included memo writing. However, because data collection was complete when coding began, we were not able to fully employ some central components of grounded theory, such as ongoing modifications to sampling and interview guides or member checking (Firmin et al., 2015; Glaser & Strauss, 1967; Heath & Cowley, 2004; Charmaz & McMullen, 2011). Coding took place in an iterative process, whereby constant comparison methods were used to develop a codebook that identified recurrent themes in the interviews; this codebook was then used to code all transcripts (Boyatzis, 1998). The qualitative software program Atlas-TI was used as a means to organize data and coding. This program does not generate codes or word categories, but rather the researcher identifies and labels themes and codes within the software, marking text within this program and using the program as a means for storing, organizing, and accessing coded text and memos.

Using the same transcripts, the narrative data was then entered into Pennebaker's Linguistic Inquiry and Word Count program (LIWC; Pennebaker et al., 2007a) by a research team member who was not involved in the thematic analysis. The LIWC examines words in a text file and matches each word to a dictionary of more than 4,500 word stems, organized into 83 categories. Analyses produce a percentage of word matches for a given category in each

narrative transcript. This program has been developed to analyze data in 10 languages and the

English version was used for the present study. The LIWC software has been used previously

with various sources of data that range from analysis of poetry and other forms of literature

(Ireland & Pennebaker, 2010; Stirman & Pennebaker, 2001), as well as with narrative data from

undergraduate students, hospitalized patients with cancer, prison inmates (Richards, Beal,

Seagal, & Pennebaker, 2000; Slatcher & Pennebaker, 2006), and in samples of individuals with

schizophrenia-spectrum disorders (Buck, Minor, Lysaker, 2015; Cohen, St. Hilaire, Aakre, &

Docherty, 2009; Junghaenel, Smyth, & Santner, 2008; Minor et al., 2015). Examples of LIWC

categories include personal pronouns, social words (e.g., group, brother, speak), negative

emotion words (e.g., frantic, uncomfortable, terrified, tough), and future words (e.g., gonna, may,

shouldn't, will). Before data was entered into the LIWC software, all interviewer prompts were

removed from transcripts so that results generated would reflect only participant language.

Finally, we compared the findings from the thematic analyses with those from the LIWC

output. Thematic analysis codes were prepped for comparison with LIWC variables by creating

dichotomous codes in SPSS for thematic analysis codes. If a particular code was present in any

part of a participants' interview, the participant was given a '1;' transcripts where the code was

absent were given a '0.' Point-biserial correlations were then run to appraise the strength of the

relationship between categories generated by the LIWC with the presence of particular thematic

analysis codes. The LIWC program generates over 83 word categories, so for the purposes of

this example, we selected categories that were among the most frequent and which were

relatively high in face-validity in order to provide a clear illustration. Prior to analyses, LIWC

categories and corresponding dictionaries for each category were reviewed by three authors

independently; after a consensus conversation, 5 pairs were agreed upon to test as potentially

having strong relationships and the remaining pairs were hypothesized to have weaker

relationships. Specifically, (1) we anticipated that some associations between thematic analysis

codes and LIWC variables would have strong, significant correlations (i.e., $r \geq .50$) when the

thematic analysis codes reflected concrete constructs, while (2) we anticipated that other pairs

would have weak to moderate correlations ($r \leq .30$), given the more subjective or abstract nature

of these codes (Cohen, 1988). The strength and direction of our hypotheses are listed in Table 1

for each association we tested.

Results

Thematic Analysis Results

Examining the narratives using thematic analysis identified several codes that emerged as

frequently discussed topics in participant narratives. A sample of these codes are listed in Table 2

and a full listing can be found elsewhere (Firmin et al., in press; Firmin et al., 2015). Codes fell

into four main categories: perceptions of treatment, self and illness, the future, and others. Codes

identified through thematic analysis were then used to generate a conceptual theory regarding the

way participants discussed aspects of their previous experiences as part of their identity, which

thereby appeared to serve as a protective factor against self-stigma (Firmin et al., in press). In

particular, participants discussed fighting or resisting stigma and the impact this had on their

hope for the future. Key themes were tied together into a proposed model which suggest that

protective factors might interact with one's response to stigma.

LIWC Results

The LIWC software generated quantitative data that indicated what percentage of a given

participants' speech fell into pre-established categories. Participants, on average, used a total of

3585 words (range of 741 to 6818). See Table 3 for descriptive statistics regarding mean

percentage of LIWC word categories, reflecting those among the most frequent with relatively high face-validity for ease of illustration. Verbs were used most frequently in our transcripts, followed by social process words (such as "babe," "aunt," and "self"). Money, religion, and friend words were used less frequently.

Results from Combining Thematic Analyses and LIWC Findings

We hypothesized that relationships would be strong between LIWC categories and thematic analyses codes that identified relatively objective constructs and this was confirmed for the majority of associations examined (see Table 4). Strong, significant relationships ranged in strength from $r = |.53-.68|$. The strongest association was found between the thematic analysis code "Anger/Aggression" and the LIWC category for negative emotion words ($r = .65, p <.001$). Also consistent with hypotheses, most relationships where codes and LIWC categories were expected to be weak to moderate, given the less discrete nature of the thematic analysis code, weaker correlations were found, ranging from $r = |.15-.33|$. The weakest association examined was between the thematic analysis code "Identifying past improvement" and the LIWC category for insight ($r = .15, p =.32$). One finding that contradicted our hypotheses was the small, nonsignificant association between "stigma" from the thematic analysis and the LIWC category "negative emotion" ($r = .12, p = .42$).

Discussion

These findings present an example of using thematic analysis and LIWC software together, including the types of comparisons and hypotheses one might make to apply these methods in tandem. Thematic analysis themes revealed participants' perceptions regarding their treatment, self and illness, the future, and view of others that seem to impact their resiliency and that may protect against internalizing stigma. LIWC analyses revealed that action-oriented and

social process language was used most often in participant narratives. As hypothesized, we found strong associations between many of the thematic analysis codes and LIWC categories when the codes reflected more objective, discrete constructs. Overall, we suggest that these findings illustrate that the LIWC has the potential to be a powerful tool for analyzing qualitative data, allowing for rapid analyses and frequencies which often would be impractical to assess without such a tool (e.g., personal pronouns). Moreover, we believe these findings illustrate the potential for the LIWC to be used with thematic analysis coding in a relatively sophisticated way which does not add substantial burden to the researcher and has the potential to draw from the strengths of each method and generate rich qualitative findings.

Cresswell (2014) notes the potential for mixed-methods designs to maximize the strengths of qualitative and quantitative analyses and minimize the weaknesses of each approach. We believe our use of these methods together strengthened the overall findings and enhanced the rigor of the analytic process in several notable ways. First, the convergence between codes developed using thematic analysis methods and the LIWC software increased our confidence in the validity of the thematic analysis codes, particularly by identifying similar frequent topics using the LIWC method and thematic analysis. A notable strength of the LIWC software is the objective method of category generation. Seeing convergence and strong associations between LIWC categories and hypothesized thematic analysis codes could serve as one potential means of a validity-check for researchers conducting more subjective coding processes like thematic analysis; for example, this may point researchers to areas of coding that were frequently discussed but perhaps overlooked by coders. A second benefit was using the LIWC software to uncover surprising relationships, or lack of expected associations (e.g., stigma and the LIWC's negative emotion words), which steered us to investigate further. In this case, upon further

analyzing the text where stigma was discussed, we discovered participants were highly resilient and thereby expressing fewer negative emotions (Firmin et al., in press).

Benchmarks of rigorous qualitative analyses have come to include validity checks (Charmaz & McMullen, 2011; Tong et al., 2007) and triangulation of qualitative data has become one well-used stratagy to enhance qualitative reliability. The present findings advance this ongoing discussion by considering the integration of thematic analyses and text-analysis software tools as a new means of triangulation that draws from the strengths of both qualitative and quantitative methods. For example, as our findings illustrate, the LIWC software quickly analyzes categories that would be tedious to manually analyze (e.g., prononoun use, emotional valence). The thematic-analysis findings, on the other hand, provided nuanced findings regarding participant discussions of a complex construct such as stigma that would have been lost had just the LIWC software been applied. Just as member-checking and consensus-based approaches to qualitative analses have been developed to help enhance the rigor of qualitative coding (Doyle, 2007), using an objective, dictionary-based computer software to provide an additional coding of qualitative data may be a relatively quick step researchers could add to further enhance the rigor of findings. The ability to rapidly identify potential areas of unexpected findings may be one additional benefit to using these methods in conjunction. Integrating text-analysis software with traditional qualitative methods may also have a role in informing new areas of research. For example, an important use of qualitative research is generating theory that is later tested using quantitative methods (Charmaz, 2014). In the present findings, unexpected associations pointed to an area of investigation that would be ripe for more nuanced thematic-analysis or investigation through follow-up interviews.

Overall, we agree with Johnson and Onwuegbuzie's (2004) conclusion from a decade ago - that researchers should be taking advantage of the multiple analytic methods available and the strengths that accompany integrating multiple approaches. However, we feel it is now important to consider the role sophisticated text-analysis software may play in enhancing the rigor of qualitative and mixed-methods analyses. This study provides a concrete example of how such combination might work. We believe that the need to consider what constitutes rigorous use of text-analysis software has never been more important given the increasing ease and availability of word-counting tools (Johnson & Onwuegbuzie, 2004; Onwuegbuzie & Leech, 2005; Pennebaker et al., 2007b).

A key strength of the current study is its combination of qualitative and quantitative methods and tools in examining the narrative data. There are several limitations to this study that future research and writing should address. These findings present a sub-set of codes and categories; for illustrative purposes, we selected categories and themes that were among the most frequently used and that had relatively high face-validity. Future studies should build on these findings and further examine appropriate application of these methods in conjunction, particularly when coding involves less discrete or objective topics. Furthermore, the nature of the thematic analysis coding in this study lended itself to creating a corresponding numeric value (i.e., 0, 1) for each code which we could then compare to the LIWC data statistically. Future work should examine alternative methods of blending these analytic methods when using additional types of qualitative data sources. The nature of our interview guide was also semi-structured, with open-ended questions and interviewers were trained to provide minimal prompting. Researchers using more directive approaches may need to determine whether speech

being analyzed by the LIWC software reflects topics spontaneously generated by participants or whether topics are introduced by the interviewers.

Overall, given the increasing use of computer-based text-analysis programs as qualitative tools, we aimed to provide an example that illustrates the integration of a newer method, the LIWC software, thematic analysis methods in a way that built on the strengths offered by each approach. We suggest that using these in conjunction with traditional qualitative methods does not significantly add burden to the researcher and has the potential to strengthen findings, enhance rigor, and potentially inform new areas for investigation.

References

Ashworth, P. (2008). Conceptual foundations of qualitative psychology. *Qualitative Psychology: A Practical Guide to Research Methods, 2*, 4-25.

Atchison, K. A. (1996). Understanding and utilizing qualitative research. *Journal of Developmental Education, 60*(8), 716-720.

Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *BMJ : British Medical Journal, 322*(7294), 1115-1117.

Berg, B. L., & Lune, H. (2004). *Qualitative research methods for the social sciences* (Vol. 5): Pearson Boston.

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research in education: An introduction to theory and methods*: ERIC.

Bordens, K. S., & Abbott, B. B. (2002). *Research design and methods: A process approach*. New York, NY: McGraw-Hill.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage.

Bringer, J. D., Johnston, L. H., & Brackenridge, C. H. (2006). Using computer-assisted qualitative data analysis software to develop a grounded theory project. *Field methods*, *18*, 245-266.

Bryman, A. (2006). Integrating quantitative and qualitative research: how is it done? *Qualitative Research, 6*(1), 97-113.

Buck, B., Minor, K. S., & Lysaker, P. H. (in press). Lexical characteristics of anticipatory and consummatory anhedonia in schizophrenia: A study of language in spontaneous life narratives. *Journal of Clinical Psychology*.

Buck, B., Minor, K. S., & Lysaker, P. H. (2015). Differential lexical correlates of social

cognition and metacognition in schizophrenia: A study of spontaneously-generated life

narratives. *Comprehensive Psychiatry, 58*, 138-145.

Buck, B., & Penn, D. L. (2015). Lexical characteristics of emotional narratives in schizophrenia:

Relationships with symptoms, functioning, and social cognition. *The Journal of Nervous

and Mental Disease*, *203*, 702-708.

Callahan, C. M., Unverzagt, F. W., Hui, S. L., Perkins, A. J., & Hendrie, H. C. (2002). Six-item

screener to identify cognitive impairment among potential subjects for clinical research.

Medical Care, 40(9), 771-781.

Chapple, A., & Rogers, A. (1998). Explicit guidelines for qualitative research: A step in the right

direction, a defence of the 'soft' option, or a form of sociological imperialism? *Family

Practice, 15*(6), 556-561.

Charmaz, K. (2014). *Constructing grounded theory*. Thousand Oaks, CA: Sage.

Charmaz, K., & McMullen, L. M. (2011). *Five ways of doing qualitative analysis:

Phenomenological psychology, grounded theory, discourse analysis, narrative research,

and intuitive inquiry*. New York, NY: Guilford Press.

Cohen, A. S., Alpert, M., Nienow, T. M., Dinzeo, T. J., & Docherty, N. M. (2008).

Computerized measurement of negative symptoms in schizophrenia. *Journal of

Psychiatric Research*, *42*, 827-836.

Cohen, A. S., St. Hilaire, A., Aakre, J. M., & Docherty, N. M. (2009). Understanding anhedonia

in schizophrenia through lexical analysis of natural speech. *Cognition and Emotion,

23*(3), 569-586.

Coolican, H. (2014). *Research methods and statistics in psychology.* New York, NY: Psychology

Press.

Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five

approaches*. Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods

approaches.* Thousand Oaks, CA: Sage Publications.

Cutcliffe, J. R. (2000). Methodological issues in grounded theory. *Journal of Advanced Nursing,

31*(6), 1476-1484.

Devers, K. J. (1999). How will we know "good" qualitative research when we see it? Beginning

the dialogue in health services research. *Health Services Research, 34*(5 Pt 2), 1153-

1188.

Dixon-Woods, M., Shaw, R. L., Agarwal, S., & Smith, J. A. (2004). The problem of appraising

qualitative research. *Quality and Safety in Health Care, 13*(3), 223-225.

Easterby-Smith, M., Golden-Biddle, K., & Locke, K. (2008). Working with pluralism:

Determining quality in qualitative research. *Organizational Research Methods, 11*(3),

419-429.

Firmin, R. L., Luther, L., Lysaker, P. H., & Salyers, M. P. (in press). Veteran identity as a

protective factor: A grounded theory comparison of perceptions of self, illness, and

treatment among Veterans and Non-Veterans with schizophrenia. *American Journal of

Psychiatric Rehabilitation.*

Firmin, R. L., Luther, L., Lysaker, P. H., & Salyers, M. P. (2015). Self-initiated helping

behaviors and recovery from severe mental illness: Implications for work, volunteerism,

and peer-support. *Psychiatric Rehabilitation Journal, 38(4)*, 336-341.

Flick, U. (2014). *An introduction to qualitative research.* Thousand Oaks, CA: Sage

      Publications.

Gingerich, S., & Mueser, K. T. (2005). Illness management and recovery. In R. E. Drake, M. R.

      Merrens & D. W. Lynde (Eds.), *Evidence-Based Mental Health Practice: A Textbook*

      (pp. 395-424). New York: Norton.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory. Strategies for Qualitative*

      Research. New York, NY: *Mill Valley: The Sociology Press*.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment

      with data saturation and variability. *Field Methods, 18*(1), 59-82.

Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed

      methods research designs in counseling psychology. *Journal of Counseling Psychology,*

      52(2), 224.

Heath, H., & Cowley, S. (2004). Developing a grounded theory approach: A comparison of

      Glaser and Strauss. *International Journal of Nursing Studies, 41*(2), 141-150.

Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: Synchrony in

      essays, correspondence, and poetry. *Journal of Personality and Social Psychology, 99*(3),

      549.

John, W. S., & Johnson, P. (2000). The pros and cons of data analysis software for qualitative

      research. *Journal of Nursing Scholarship*, *32*, 393-397.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm

      whose time has come. *Educational Researcher, 33*(7), 14-26.

Junghaenel, D. U., Smyth, J. M., & Santner, L. (2008). Linguistic dimensions of

psychopathology: A quantitative analysis. *Journal of Social and Clinical Psychology,*

27(1), 36-55.

Kuper, A., Lingard, L., & Levinson, W. (2008). Critically appraising qualitative research. *British*

Medical Journal, 33(7), 687-692.

La Pelle, N. (2004). Simplifying qualitative data analysis using general purpose software tools.

Field Methods, 16(1), 85-108.

Lee, C. H., Lee, M., Ahn, S., & Kim, K. (2007). Preliminary analysis of language styles in a

sample of schizophrenics. *Psychological Reports, 101*, 392-394.

Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: A call for

data anlysis triangulation. *School Psychology Quarterly, 22,* 557-584.

Leech, N. L., & Onwuegbuzie, A. J. (2009). A typology of mixed methods research designs.

Quality and Quantity, 43(2), 265-275.

Leiva, F. M., Ríos, F. J. M., & Martínez, T. L. (2006). Assessment of interjudge reliability in the

open-ended questions coding process. *Quality and Quantity, 40*(4), 519-537.

Lysaker, P. H., Carcione, A., Dimaggio, G., Johannesen, J., Nicolò, G., Procacci, M., &

Semerari, A. (2005). Metacognition amidst narratives of self and illness in schizophrenia:

Associations with neurocognition, symptoms, insight and quality of life. *Acta*

Psychiatrica Scandinavica, 112(1), 64-71.

Lysaker, P. H., Clements, C. A., Plascak-Hallberg, C. D., Knipscheer, S. J., & Wright, D. E.

(2002). Insight and personal narratives of illness in schizophrenia. *Psychiatry:*

Interpersonal and Biological Processes, 65(3), 197-206. Lowe, W. (2002). Software for

content analysis–A Review. *Cambridge: Weatherhead Center for International Affairs and the Harvard Identity Project*.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (Vol. 1). New York, NY: Psychology Press.

Mayoh, J., & Onwuegbuzie, A. J. (2015). Toward a conceptualization of mixed methods phenomenological research. *Journal of Mixed Methods Research, 9*(1), 91-107.

Mays, N., & Pope, C. (1995). Rigour and qualitative research. *British Medical Journal, 311*(6997), 109-112.

Mays, N., & Pope, C. (2000). Assessing quality in qualitative research. *British Medical Journal, 320*(7226), 50-52.

Meyer, D. Z., & Avery, L. M. (2009). Excel as a qualitative data analysis tool. *Field Methods, 21*(1), 91-112.

Minor, K. S., Bonfils, K. A., Luther, L., Firmin, R. L., Kukla, M., MacLain, V. R., … & Salyers, M. P. (2015). Lexical analysis in schizophrenia: How emotion and social word use informs our understanding of clinical presentation. *Journal of Psychiatric Research, 64*, 74-78.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative Health Research, 8*(3), 362-376.

Morse, J. M. (1997). " Perfectly healthy, but dead": The myth of inter-rater reliability. *Qualitative Health Research, 7*(4), 445-447.

Onwuegbuzie, A. J., & Leech, N. L. (2005). On becoming a pragmatic researcher: The importance of combining quantitative and qualitative research methodologies. *International Journal of Social Research Methodology, 8*(5), 375-387.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007a). Linguistic Inquiry and Word Count: LIWC2007. Austin, TX, United States: LIWC.net.

Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007b). The development and psychometric properties of LIWC2007. Austin, TX, United States: LIWC.net.

Pope, C., & Mays, N. (1995). Qualitative research: Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *British Medical Journal, 311*(6996), 42-45.

Pratt, M. G. (2008). Fitting oval pegs into round holes: Tensions in evaluating and publishing qualitative research in top-tier North American journals. *Organizational Research Methods, 11*(3), 481-509.

Pratt, M. G. (2009). From the editors: For the lack of a boilerplate: Tips on writing up (and reviewing) qualitative research. *Academy of Management Journal, 52*(5), 856-862.

Richards, J. M., Beal, W. E., Seagal, J. D., & Pennebaker, J. W. (2000). Effects of disclosure of traumatic events on illness behavior among psychiatric prison inmates. *Journal of Abnormal Psychology, 109*(1), 156.

Ritchie, J., Lewis, J., Nicholls, C. M., & Ormston, R. (2013). *Qualitative research practice: A guide for social science students and researchers.* Thousand Oaks, CA: Sage Publications.

Sale, J. M., Lohfeld, L., & Brazil, K. (2002). Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality and Quantity, 36*(1), 43-53.

Salyers, M. P., Matthias, M. S., Sidenbender, S., & Green, A. (2013). Patient activation in
schizophrenia: Insights from stories of illness and recovery. *Administration and Policy in
Mental Health and Mental Health Services Research, 40*(5), 419-427.

Sandelowski, M. (2000). Focus on research methods-whatever happened to qualitative
description? *Research in Nursing and Health, 23*(4), 334-340.

Sandelowski, M. (2001). Real qualitative researchers do not count: The use of numbers in
qualitative research. *Research in Nursing & Health, 24*(3), 230-240.

Shah, S. K., & Corley, K. G. (2006). Building Better Theory by Bridging the Quantitative–
Qualitative Divide. *Journal of Management Studies, 43*(8), 1821-1835.

Slatcher, R. B., & Pennebaker, J. W. (2006). How do I love thee? Let me count the words: The
social effects of expressive writing. *Psychological Science, 17*(8), 660-664.

Smith, J. A. (2007). *Qualitative psychology: A practical guide to research methods.* Thousand
Oaks, CA: Sage Publicatinos.

St. Hilaire, A., Cohen, A. S., & Docherty, N. M. (2008). Emotion word use in the conversational
speech of schizophrenia patients. *Cognitive Neuropsychiatry, 13*(4), 343-356.

Stirman, S. W., & Pennebaker, J. W. (2001). Word use in the poetry of suicidal and nonsuicidal
poets. *Psychosomatic Medicine, 63*(4), 517-522.

Suddaby, R. (2006). From the editors: What grounded theory is not. *Academy of Management
Journal, 49*(4), 633-642.

Swallow, V., Newton, J., & Van Lottum, C. (2003). How to manage and display qualitative data
using 'Framework'and Microsoft® Excel. *Journal of Clinical Nursing, 12*(4), 610-612.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and

    computerized text analysis methods. *Journal of Language and Social Psychology, 29*(1),

    24-54.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative

    research (COREQ): A 32-item checklist for interviews and focus groups. *International

    Journal for Quality in Health Care, 19*(6), 349-357.

Weiss, R. S. (1995). *Learning from strangers: The art and method of qualitative interview

    studies*: Simon and Schuster.

Table 1. Hypothesized Relationships Between the LIWC Software and Thematic Analysis Codes

Thematic Analysis Method	LIWC Software Method	Hypothesized Direction
Hypothesized Strong Relationships		
Anger/Aggression	Negative emotion	+
Friendships/Social support	Friend	+
Money issues	Money	+
Religion/Spirituality	Religion	+
Stigma	Negative emotion	+
Hypothesized Weak to Moderate Relationships		
Helping others	Social process	+
Identifying past improvement	Insight	+
Isolation	We	-
Life goals	Future	+
Medication	Health	+
Optimism: Belief future will improve	Tentative	-
Recovery perspective	Positive emotion	+
Relapse	Achievement	-
Suicide	Positive emotion	-
Want to hide symptoms	Social process	+
Wronged by others	Negative emotion	+

Table 2. Sample Codes Identified Using Thematic Analysis Analyses

Category	Code
Perceptions of Treatment	Medication
	Mental health treatment
	Recovery perspective
	Relapse
Perceptions of Self and Illness	Anger/Aggression
	Isolation
	Stigma
	Suicide
	Want to be normal/Hide symptoms
Perceptions of the Future	Identifying past improvement
	Life goals
	Money issues
	Optimism: Belief the future will improve
Perceptions of Others	Friendships/Social support
	Helping others
	Religion/Spirituality
	Wronged by others

Table 3 – Descriptive Statistics for Selected LIWC Categories

LIWC Category	Mean	Standard Deviation
Negative emotions	1.62	0.53
Friend	0.17	0.12
Health	1.32	0.64
Money	0.33	0.26
Religion	0.22	0.19
Social Process	10.15	2.59
Insight	3.57	1.62
We	0.48	0.32
Future	0.73	0.41
Tentative	2.75	1.02
Verb	20.34	2.35
Achievement	1.25	0.45
Positive emotions	2.37	0.67

Note. Mean values represent the average percentage of total words in a transcript which belong to a given LIWC category.

Table 4. Codes Identified Using Thematic Analysis compared to the LIWC Software

Thematic Analysis Method	LIWC Software Method	Point Biserial Correlation
Hypothesized Strong Relationships		
Anger/Aggression	Negative emotion	.68***
Friendship/Social support	Friend	.42***
Money issues	Money	.63***
Religion/Spirituality	Religion	.53***
Stigma	Negative emotion	.12
Hypothesized Weak to Moderate Relationships		
Helping others	Social process	.26
Identifying past improvement	Insight	.15
Isolation	"We"	-.32*
Life goals	Future	.31*
Medication	Health	.31*
Optimism: Belief future will improve	Tentative	-.33*
Recovery perspective	Positive emotion	.18
Relapse	Achievement	-.18
Suicide	Positive emotion	-.30*
Want to be normal/Hide symptoms	Social processes	.30*
Wronged by others	Negative emotion	.24

* Statistical significance at p<.05, ** Indicates statistical significance at p<.01, *** Indicates statistical significance at p<.001.