Mixed methods research Part 1: Collecting and analysing social media textual data using netnography and thematic analysis

Denitsa Dineva, Cardiff University

## Abstract

This dataset article explains the first stage of a mixed methods research design set within an overall qualitative piece of research. It provides a step-by-step guide on accessing, recording, and analysing qualitative data via the methods of netnography and thematic analysis. Netnography is a common method for researching online spaces such as social media pages and online forums to generate qualitative insights about a topic of interest. Thematic analysis, in turn, translates these insights into distinct themes that reflect the key findings in the data. This dataset article refers to an Excel (.xlsx) file that contains social media textual data collected to understand what conflict management strategies a brand uses to moderate uncivil consumer comments on its Facebook page.

Student Guide

Introduction

Qualitative research involves the collection of non-numerical data to understand social phenomena such as people's beliefs, experiences, attitudes, behaviours, and interactions. As such, qualitative research methods focus on exploring a research problem that is not well understood, but often do not provide evidence on causation, that is the relationship between the cause and effect of the studied social phenomenon. In turn, mixed methods research combines qualitative research techniques with quantitative ones in order to provide a thorough understanding of a specific social phenomenon. Among the most common approaches to conducting mixed methods research is a sequential exploratory design whereby qualitative research techniques precede and underpin a quantitative research stage (Figure 1) (Cresswell et al., 2003). This article explores the qualitative stage of mixed methods research with a particular focus on *netnography* as a data collection tool together with *thematic analysis* as a data analysis tool.
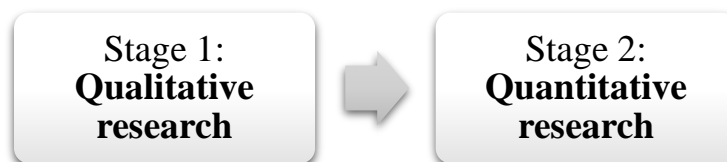
Figure 1. Sequential exploratory design

This article dataset is taken from Dineva et al.'s (2020) research paper, which explores the different strategies a brand uses to moderate uncivil consumer comments. This is done by systematically recording relevant social media textual data using netnographic observations and analysing these data via data-driven thematic analysis.

Netnography and Thematic Analysis

Netnography is a well-established form of qualitative research. The method adopts the principles behind ethnography, which consists of employing in-person/offline observations of social phenomena (e.g., people, behaviours) and their environments. In contrast, netnography aims to understand social phenomena in contemporary digital/online contexts. Netnography involves conducting observations of computer-mediated environments (e.g., chatrooms, social media, online forums) in relation to topics of interest and collecting data correspondingly (Kozinets, 2002). For example, a researcher may be interested in gaining an insight into consumers' opinions regarding a new product introduced by a brand on the market. Netnographic observations of the brand's posts and consumer comments on social media can be valuable to gain insights into initial consumer attitudes and behaviours towards the product. Compared with other qualitative methods such as face-to-face focus groups, interviews, and ethnographic research, netnography is less time-consuming, more flexible and generates insights that are dependable, despite its comparative ease of use (Morais et al., 2020). Netnography has been applied in a variety of ways across different social sciences fields and topics including consumer decision-making, business-to-business, service ecosystems, and sustainability among others, which are discussed in detail in Heinonen and Medberg's (2018) work.

While netnography is a research tool adaptable to studying a wide range of research topics pertinent to online spaces (e.g., social media), thematic analysis represents a foundational method for the analysis of already collected qualitative data (Braun and Clarke, 2006) that netnography has facilitated. Thematic analysis is a method for identifying, analysing, and reporting patterns (themes) within data and its main advantage is flexibility with research topics, environments, and paradigms. At the minimum, it organises and describes a dataset in (rich) detail, but it can go further than this to interpret various aspects of a research

topic. For example, the researcher may be interested in delineating between different categories (themes) of consumer sentiment (e.g., positive, negative, neutral) in relation to healthy eating or he/she may wish to explore deeper questions (themes) such as the underlying issues including body image, self-esteem and eating disorders.

In sum, netnography and thematic analysis are two qualitative research methods that can often go hand in hand due to their wide applicability and flexibility to suit different research paradigms and topics.

Data excerpt

The data excerpt consists of an .xlsx file and requires access to Microsoft Office and Excel specifically. The data comprise of textual data excerpts (brand posts in response to uncivil consumer comments) and these have been taken from Dineva et al. (2020) who draw data from the official Facebook page of PETA ("People for the Ethical Treatment of Animals"). To protect the privacy of the individuals, real names have been hidden in excerpts that contain uncivil consumer comments, which is a standard practice in qualitative research.

Data Collection and Analysis

Data collection and analysis largely consist of three main stages: (1) selecting appropriate research sites and collecting relevant data, (2) analysing the data to develop stand-alone and coherent themes, and (3) reporting the final results. These are explained in detail here.

Stage 1: Site selection and data collection

The first step involves the researcher defining her/his research questions and then identifying relevant online spaces, groups, or communities that are deemed appropriate in answering these research questions. In the present dataset, the researchers were interested in

uncovering how brands respond to uncivil consumer interactions on their social media pages. To answer this research question, the authors identified the non-profit brand PETA and its official Facebook page as an appropriate site for data exploration and collection via the method of netnography.

Prior to collecting data, several ethical and research considerations have to be made regarding the type of online space used and any access restrictions as well as participant consent and privacy. First, determining whether the data required are located in an open or private forum is a key consideration for researchers. As a rule of thumb, if the group is closed (requires registration or admin approval before joining), the researcher must disclose their presence in the group and obtain approval from its administrator or members before collecting any data. In other words, if the group is closed, the researcher must seek consent from the administrator to collect data. Second, when collecting data, the researcher has the choice between disclosing their presence (overt research) or not (covert research). This is dependent not only on the type of online space being accessed, as already discussed, but also on the research topic and whether this requires the researcher's active participation in comments and posts. Third, when collecting data, the researcher must either remove any personal information that identifies an individual (e.g., name) or use pseudonyms to cover their identity. In the present dataset, the researchers accessed a public page (open to non-members) and, therefore, did not disclose their presence or participate in the discussions (Langer and Beckman, 2005).

Once the researcher has identified suitable online groups and/or communities, the researcher must ensure they meet the following criteria to justify their selection for netnography:

- A focus on a specific social phenomenon relevant to the research question,
- High "traffic" of postings, and

- Availability of relevant to the research topic data.

After the researcher has chosen the online space(/s), he/she can begin data collection, which consists of two elements: the data the researcher directly copies or screenshots from the online site (verbatim quotes) and the data the researcher inscribes regarding his/her observations (informal notes for future reference). In this dataset, the researchers collected excerpts from PETA's official Facebook page consisting of brand conflict management strategies used to address uncivil consumer comments.

Stage 2: Data analysis

The second stage is data analysis. This involves thematic analysis, which is a method that systematically identifies, analyses and reports patterns of meaning (themes) across a set of qualitative data (Braun and Clarke, 2014). Thematic analysis involves six stages, as outlined below.

1. *Familiarization with the data*: Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas.

2. *Generating initial codes*: Coding interesting features of the data in a systematic fashion across the entire dataset.

3. *Searching for themes*: Collating codes into potential themes, generating all data relevant to each potential theme.

4. *Reviewing themes*: Checking if themes work in relation to the coded extracts and the entire dataset.

5. *Defining and naming themes*: Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.

6. *Reporting the results*.

In *Phase 1*, the researcher copies relevant data from the online space to a Word document. In their dataset, the authors copied 343 screenshots of relevant data excerpts from PETA's Facebook page into a Word document and removed the names of the participants to protect their identity, as shown below in Excerpts 1&2.

Data excerpt #1

Data excerpt #2

This document represents the raw data that are kept by the researcher as a record before these have been further processed, cleaned, or analysed. It is important to keep a copy of the pre-processed data (e.g., screenshots), because this allows the researcher to go back to the data and get further insights into the data or for clarification purposes. The researcher must then repeatedly read the data excerpts to familiarise himself/herself with the data and in order to begin searching for patterns and re-occurring meanings in relation to the research questions. In this dataset, the authors were interested in the brand posts alone and searching for patterns within these. For ease of analysis, once the researcher has identified their focal point of

analysis, these data alone can be transferred to an Excel spreadsheet in a similar manner to the dataset supplied here, which contains only the brand posts.

During *Phase 2* of the thematic analysis, the researcher can begin to generate codes that correspond to a feature of the data and in this case an element of brand conflict management. An important decision that needs to be made here is whether the codes are derived from the data alone and thus are strongly linked to the data themselves, or from pre-existing theory where codes are pre-defined and applied to the data or a mix of both (a hybrid approach). You can read more about the different approaches to coding in Braun and Clark (2014) and Fereday and Muir-Cochrane (2006). In the present dataset, the authors chose to focus on generating codes from the data alone (i.e., *data-driven coding*). Indeed, thematic analysis is usually employed as a technique that is driven by the data and the repeated readings of these generate theoretical concepts.

In *Phase 3*, the researcher begins to analyse the codes and consider the ways in which these can be combined to construct an overarching theme (i.e., a candidate conflict management strategy). These two phases are exemplified in the following two coding examples (Excerpts 3&4). The researchers first noted that the two data excerpts showcase a similar pattern, that is the brand providing an explanation to uncivil consumers and this is evidenced in the similarity between the codes assigned to the two excerpts. Following several occurrences of these codes, the authors began to observe a pattern between these brand strategies, which constitutes a candidate theme.

making money has everything to do with it. When people are financially motivated the interest of others is never a priority. Additionally, comparing an animal that has been systematically bred to rely on another species to live, to that of an animal that has evolved to strive in their social pack, is a weak misguided comparison.

Like · Reply · 3 · 12 August at 14:45 · Edited

Zoo's have been established originally by taking wild animals from their habitat, away from their families, and putting them in cages in another climate to exhibit and profit from it. Humans idea, it's wrong!!

Like · Reply · 4 · 12 August at 14:43

PETA (People for the Ethical Treatment of Animals) ● Zoos all over the U.S. have closed their elephant exhibits or announced that they intend to phase them out, citing an inability to provide the animals with proper care. There is absolutely no ethical way to keep these intelligent, social animals in captivity.

Like · Reply · 6 · 12 August at 14:43

```
provides an explanation
to an issue causing the
conflict
explanation possibly
aimed
at educating
conflicting
parties
aligned with company
mission
```

Coding data example #3

Someone obviously didn't read.

Like · Reply · 6 August at 02:19

PETA (People for the Ethical Treatment of Animals) ● We are recommending release into coastal sanctuaries, not directly into the wild. http://www.seaworldofhurt.com/sea-sanctuaries-sure-thing.../

Sea Sanctuaries Are a Sure Thing— Get With the Program, SeaWorld -...

SEAWORLDOFHURT.COM

Like · Reply · 6 August at 18:05

```
Further explains an issue
that is causing the
conflict
Aligned with company
mission
Provides hyperlink
```

Coding data example #4

These codes were subsequently subsumed into a distinctive conflict management strategy based on unifying features and similarity between the codes – in this case a strategy labelled as "educating". The researchers followed this process – from coding, to identifying recurring patterns between similar codes, to developing themes until all data were coded into distinctive themes. Data that were analysed using codes, which occurred highly infrequently or did not match any other codes, should be removed from further analysis, and excluded from the findings.

In *Phase 4*, a review of the set of constructed themes is conducted in order to refine the themes. The refinement involves two main criteria: *internal homogeneity* and *external heterogeneity*. In other words, the researcher aims to ensure that data within the constructed themes are similar enough (internal homogeneity), while clear and identifiable distinctions between the themes exist and these do not overlap significantly (external heterogeneity). During this phase, a researcher may find that data excerpts that are assigned the same/similar code are too diverse, making it challenging to produce a coherent theme and must thus be excluded from further analysis. In the next phase (*Phase 5*), the themes are finalised by developing definitions and naming each theme (conflict management strategies).

A brief note on data analysis. This dataset discusses manual data analysis, but automated approaches to qualitative coding such as NVivo are also popular and widely available. The choice between manual versus automated data analysis lies with the researcher and their goal(/s) as well as the volume of data available for analysis.

Stage 3: Producing a results report

The final step involves producing the written report of the analysed data in a coherent and systematic manner. In this phase, the researcher must select suitable and vivid data extracts to showcase the distinct themes (brand conflict management strategies) identified in the data. It is good practice for the reports to present the codes underpinning each theme, the final themes, and their frequency of occurrence (in numbers or percentage), definitions of the themes, and examples from the dataset to help illustrate each theme, as demonstrated in Table 1.

Codes	Themes	No of observations	Definition	Data excerpt
Not intervening in the conflict Avoiding engaging in the conflict	Non-engaging	265	The organization does not take any action to moderate a conflict.	
Missing comments	Censoring	2	The organization permanently removes consumer comments.	*[comment was removed]*
Thanks cause supporter(s) Agrees with cause supporter(s)	Bolstering	14	The organization affirms a consumer comment.	*"Thank you for choosing compassion! (heart emoji) #FriendsNotFood #TheYearOfVegan"* *"@Lisa thanks for explaining supply & demand. (winking face emoji)"*
Further clarification about an issue causing the conflict Explaining an issue to conflicting parties Providing additional information about an issue	Educating	21	The organization provides educational information about an ethical issue.	*"Zoos claim to provide educational opportunities, but most visitors spend only a few minutes at each display, seeking entertainment rather than enlightment [sic]."* *"Keeping animals in cages does nothing to foster respect for animals since all children learn is that animals will spend their lives behind bars for people's fleeting distraction and amusement."*
An appeal to take action	Mobilizing	30	The organization urges consumers to take action towards an ethical issue.	*"Unfortunately, a majority of dairy farms use practices like the ones seen in this video. Please consider ditching dairy and going vegan:*

| Urging conflicting parties to change their behavior | | | | *http://features.peta.org/how-to-go-vegan/"*

"Please tell everyone you know to go vegan to help stop this! http://www.peta.org/living/food/free-vegan-starter-kit/" |

Table 1 Sample findings report

Summary

In summary, this dataset article demonstrated how netnography and thematic analysis can be used to generate qualitative research insights into real brand practice (conflict management strategies). The article discussed netnography as a suitable method for collecting social media data and the steps involved in conducting netnographic observations. In terms of qualitative data analysis, this article explained how thematic analysis can be utilised, specifically focusing on manual data analysis, and illustrating its execution step-by-step.

To find out more about how this qualitative research can be followed up with quantitative research, refer to dataset article: *"Mixed methods research Part 2: Designing an online experimental survey from prior qualitative research and using one-way ANOVA in SPSS for data analysis"*.

Reflective questions

Can you identify a research problem that netnography can help you understand better?

Which approach to thematic analysis will suit your data better: theory- or data-driven?

What ethical issues could arise from researching your chosen research topic and what key steps will you undertake to address these?

Looking at the dataset in the Excel spreadsheet, how does your analysis of the data compare with the analysis provided? Are there any different or other levels of analysis you managed to identify?

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. Doi: 10.1191/1478088706qp063oa

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of Mixed Methods in Social and Behavioral Research*, *209*(240), 209-240.

Dineva, D., Breitsohl, J., Garrod, B., & Megicks, P. (2020). Consumer responses to conflict-management strategies on non-profit social media fan pages. *Journal of Interactive Marketing*, *52*, 118-136. Doi: 10.1016/j.intmar.2020.05.002

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods, 5*(1), 80-92. Doi: 10.1177/160940690600500107

Kozinets, R. V. (2002). The field behind the screen: Using netnography for marketing research in online communities. *Journal of Marketing Research*, *39*(1), 61-72. Doi: 10.1509/jmkr.39.1.61.18935

Langer, R., & Beckman, S. C. (2005). Sensitive research topics: netnography revisited. *Qualitative Market Research: An International Journal, 8*(2), 189- 203. Doi: 10.1108/13522750510592454

Morais, G. M., Santos, V. F., & Gonçalves, C. A. (2020). Netnography: Origins, foundations, evolution and axiological and methodological developments and trends. *The Qualitative Report*, *25*(2), 441-455. Doi: 10.46743/2160-3715/2020.4227

Heinonen, K. and Medberg, G. (2018), Netnography as a tool for understanding customers: implications for service research and practice, *Journal of Services Marketing*, *32*(6), 657-679. Doi: 10.1108/JSM-08-2017-0294

Further reading

Dineva, D., Breitsohl, J. C., & Garrod, B. (2017). Corporate conflict management on social media brand fan pages. *Journal of Marketing Management*, *33*(9-10), 679-698. Doi: 10.1080/0267257X.2017.1329225

Coolican, H. (2017). *Research methods and statistics in psychology*. Psychology Press.

Feigenbaum, C. (2018). Netnography: Capturing Consumer Behavior, Available at: https://netbasequid.com/blog/netnography-2017-capturing-consumer-behavior/

Kozinets, R. V. (2006). Netnography. *Handbook of Qualitative Research Methods in Marketing*, 129-142.

Morse, J. M. (2016). *Mixed method design: Principles and procedures*. Routledge.

Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. *The SAGE Handbook of Qualitative Research in Psychology*, *2*, 17-37.