

V1P1C4b. Qualitative insights for digital marketing

Qualitative research has often been referred to as “a craft” (Seale 1999; Kvale and Brinkmann 2009), in reference to how it requires the researcher’s involvement in creating by hand (i.e. not in an automated way) an inquiry and generating insights about a phenomenon. As craft, qualitative research is not always fast, precise, or accurate. Craft requires patience, attention to detail, and a good deal of creativity. In fact, as Denzin and Lincoln (2017) note, the qualitative researcher can be seen as a *bricoleur*—someone who is able to work with whatever resources (strategies, tools, methods, data sources) they have at hand to produce a solution tailored to a specific need. The accelerated and kaleidoscopic nature of digital environments may seem at odds with bricolage and craft-like approaches to researching it.

In this chapter, we argue that qualitative research can be mobilized to generate insights on digital marketing when researchers adopt a dynamic methodological practice and guide their qualitative studies by the flexible principles of thematic analysis. In the next sections, we explain how these aspects manifest in research projects by introducing the notion of research scenarios. We then discuss qualitative research analytical process and explain the coding procedure. Lastly, we introduce thematic analysis as an analytical approach that is suitable to the flexible research scenarios of digital research in marketing. To illustrate these discussions, we leverage a single example of a research project conducted by a team of marketing and management researchers who set out to examine how consumer practices are disrupted by governmental initiatives to implement sustainable consumption measures. The research report that resulted from the project was published by Gonzalez-Arcos et al.¹ (2021) as a theory of consumer resistance to

¹ One of the authors of this chapter, Daiane Scaraboto, is a coauthor in this article. Not all information about data collection and analysis shared here (with the consent of coauthors) has been included in the published article. When that is the case, we refer to the information as Scaraboto (n.d.).

sustainability interventions.

Digital qualitative research

Digital hyperconnectivity produces an abundance of data that makes the life of digital marketing researchers exciting and challenging in equal ways. The popularization of smartphones (Reed 2018), the (somewhat) affordable access to Internet (Plant 2004) and the widespread use of digital technologies has blurred distinctions between people's virtual and corporeal existence. So much so that we can no longer separate people's actions into two domains; online or offline behavior. Much less can we say that feelings and opinions do not transfer across these two domains, as people can no longer leave behind their online existence when they are "just" existing in the corporeal world. Furthermore, millions of people enact more than one online persona, which multiplies their experiences and practices, and associated feelings and opinions.

The research community responded to these challenges by developing or adapting research methods in line with their research philosophy. Research philosophies are grounded on and differentiated by a set of research assumptions. The most common research philosophies in management, business, and, marketing navigate between the objectivism–subjectivism continua (Saunders, Lewis and Thornhill 2019, 159):

- "Objectivism incorporates assumptions of the natural sciences. It entails realist ontology (which holds that social entities exist in reality external to and independent from social actors), epistemology focused on the discovery of truth by means of observable, measurable facts, and claims to have a value-free, detached axiology."
- "Subjectivism incorporates assumptions of the arts and humanities. It entails nominalist ontology (which holds that social phenomena are created through the language, perceptions and consequent actions of social actors), epistemology focused on the social actors' opinions, narratives, interpretations, perceptions that convey these social realities, and claims to have a value-bound, reflexive axiology." (Saunders, Lewis and Thornhill 2019, 159)

Digital marketing researchers complying with *Objectivism* have translated conventional quantitative research practice to online environments (e.g., Brügger and Willems 2009) and more recently embraced the big data approach (see also: [link to chapter in the handbook](#)). Big data analytics came as a solution to the fact that the speed and quantity of data being generated online seemed to require automated “power” (i.e. algorithms, artificial intelligence) to effectively collect, organize, and categorize data for analysis. Digital marketing researchers aligned with *Subjectivism* have either updated well-established qualitative research practices by introducing new data collection procedures (i.e. online ethnography) or developed new ones that were suited to new online environments, such as netnography (Kozinets, 2019; see also: [link to chapter in the handbook](#)).

Contrary to common understanding, qualitative studies are not research practices working in opposition to quantitative ones. Rather, qualitative and quantitative are essentially research practices that accomplish similar goals but in distinct ways (Denzin and Lincoln, 2008). This happens because each practice complies to a distinct research philosophy which, in turn, differs on epistemological principles. As a result, different data types will be prioritized in qualitative or quantitative studies: numerical data prevails in quantitative studies, whilst verbal, textual and visual are given preference in qualitative studies datasets. Textual data is collected from interviews, observations, or writing published by others, whereas verbal data is collected from speech (i.e. people’s spoken words), and visual data can be found, or created, in a variety of formats such as drawings, photographs, and video (Saunders, Lewis and Thornhill 2019). In this chapter, we focus on how the various forms of digital qualitative data can be analyzed to generate insights for marketing.

Qualitative research studies have been dealing with a diversity of verbal, textual and visual

for more than a century (Silverman 2013), and the data collected by qualitative researchers who do research in digital contexts is not new in this regard. Essentially, qualitative digital datasets will have a combination of these three data formats, even if encompassing with user-generated, computer-mediated content. As a result, the nature of qualitative inquiry does not change dramatically in transitioning from offline to online domains.

The qualitative researcher will be working with data in-context and going beyond the first, literal layer of meaning in text, verbal and visual data, to search for in-depth meaning in conversation, discourse, and narrative. In digital qualitative research, meanings are also constructed through similar a analytic process. That is, through the identification of patterns (coding), themes, and the researcher's close engagement with the data. Therefore, qualitative digital datasets will mainly differ from non-digital ones in terms of volume, and format richness (including moving images, symbols and hashtags). As such, digital qualitative researchers will also aim to develop interpretations (rather than descriptions) of their datasets, and generate insights (rather than readings) as an outcome. For this reason, research programs can transmute to the digital environment without sacrificing their research principles and practices (e.g. Caterall and Maclaran, 2002; Eagar and Dann 2016).

Research scenarios

Research scenarios are the combination of the phenomenon of interest, researchers' position in relation to the data, and the interpretive lens that will inform the development of insights in a research project. This "research assemblage" (Fox and Alldred 2015) captures the essential parts that make a research project unique, and that, when changing, create a new scenario, transforming the project. When it comes to qualitative research for digital marketing, research scenarios can vary widely across projects and change multiple times through the duration of a

single research project. We offer a brief overview of the essential elements of research scenarios (see, also, Figure 1) and their variation in order to highlight the importance of adopting a flexible methodological approach in qualitative digital research, including an analytical technique that is suitable to multiple and changing research scenarios.

The phenomenon of interest is the starting point of many qualitative inquiries. With phenomena that are digital or that manifest in digital media, digital affordances immediately become relevant and assume a fundamental role in research scenarios. Most digital phenomena are profoundly shaped by the platforms in which they unfold. Take, for example, the research project developed by Gonzalez-Arcos et al. (2021) where the researchers set out to investigate how consumer practices were affected by sustainability interventions. When the research team began the project, a nation-wide ban on plastic bags had just been implemented in Chile, and the researchers wanted to capture how consumers reacted to the ban in that context, in real time. Although the phenomenon had offline manifestations (e.g. with consumers complaining at checkout points in supermarkets), most consumers manifested their reactions online, in Twitter threads, Facebook posts, and comments on news articles on media websites. As the digital aspects of the phenomenon were prevalent, the researchers started their data collection on Twitter, because consumers replied to posts made by the Chilean government to promote the ban and hijacked the governmental campaign's official hashtag (#chaobolsasplasticas = #byeplasticbags) to manifest their support to- and discontent with the intervention.

The second element of the research scenario, researchers' capacities and position in relation to the data are a limiting and/or potentializing condition for the scope of the data collection. As Lester and Paulus (2021, 15) note, "researchers in the digital world are not only collaborating more often with each other, but, increasingly, with participants and citizen-researchers [...] part

of this collaboration is between the researchers and the material artifacts (digital tools) they employ.” Researchers who are more versus less familiar with certain digital platforms, analytical software, and data formats, will find themselves in different research scenarios—perhaps favoring one type of data over another, or struggling to understand participants’ practices. Through time, a researcher may acquire skills or develop familiarity with certain platforms, and this will change the scenario, allowing for additional types of data to be collected, deeper engagement with research participants, etc.

In the case of Gonzalez-Arcos et al. (2021)’s project, reflecting on the capacities of the research team at that time led the authors to collect data in multiple formats: qualitative data (text and images from Twitter and Facebook posts) and visual network data using a network analysis application, to facilitate the identification the key participants in this conversation (i.e. top Twitter accounts), and to provide a visual overview of the networks being developed. The visual network data allowed the researchers to identify other hashtags consumers were using in combination with #ByePlasticBags and collect data on those hashtags. These data were complemented with textual data including media articles retrieved from news media websites, and regulatory documents tracing the history of the development of the sustainability intervention.

Finally, in academic research projects—theories previously associated with a phenomenon of interest provide the ontological and epistemological basis for the study. Theoretical frameworks shape the digital methodology by setting boundaries around a phenomenon, and by assigning ontological and epistemological approaches to the project. A selection of such theories (or concepts and principles that are part of these theories) will guide researchers’ attention to specific sources of data and data types, also informing their interpretation of the digital datasets

for generating insights.

To keep with our working example, Gonzalez-Arcos and her colleagues explored some theoretical frameworks (e.g. institutional theory, normalization process theory) through the development of their project. At some point during initial rounds of data collection and analysis, the researchers settled on social practice theories (Schatzki 2018; Shove, Pantzar, and Watson 2012) as the theoretical lens informing their analysis and interpretation of the phenomenon. With that, network data became less relevant, as it was incompatible with the ontology and epistemology of the theoretical framework. Images shared by consumers, on the other hand, gained importance as a source of insight about how consumers performed the shopping practice without disposable plastic bags. Ultimately, the research design was determined by a combination of these three elements of the flexible digital methodology: the evolving phenomenon, the capacities of the research team (which may also change with the addition of new members to the research team or the acquisition of new methodological/analytical skills by existing members), and the theoretical framework being employed.

Understanding that phenomenon, researcher capacities and theoretical frameworks influence one another and require mutual adjustments is important to be able to design flexible digital research methodologies, and to adjust the methodological approach for the generation of marketing insights throughout the unfolding of a research project.

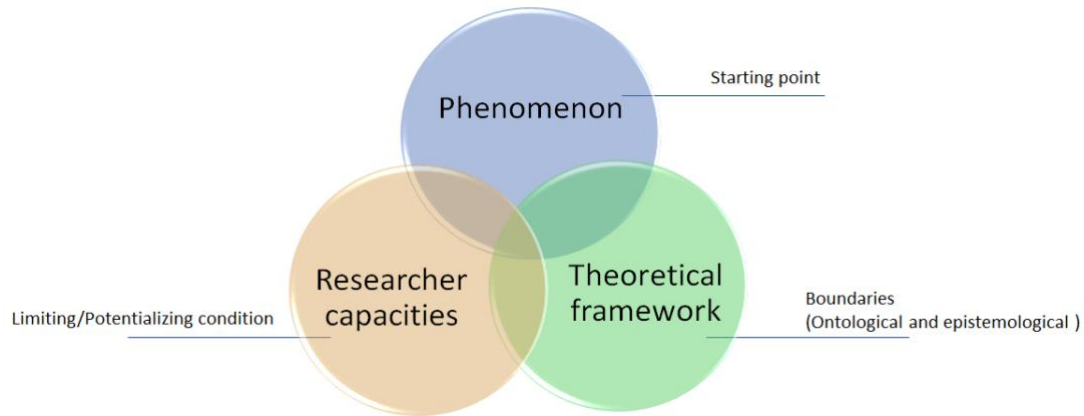


Figure 1 – Research Scenarios (Source: Visual representation created by the authors)

Thematic Analysis

Thematic analysis is a method that supports the researcher in finding patterns in the data. Rather than a theoretically embedded research program, *thematic analysis* offers a theoretically flexible analytic approach that can be adopted in a variety of qualitative studies; phenomenological examination, everyday subjective experience or critical interrogation (Braun and Clark 2012; 2019). This method allows for the emergence of insights from comparing multiple sources of data, which makes it suitable to research in digital contexts, where consumers navigate from one platform to another seamlessly, and data collection is usually designed to capture the complexity of meaning that emerges from the flexible digital landscape. By identifying patterns and—most importantly—meaningful ones, researchers can develop their empirically-based interpretation of a cultural phenomenon, and generate insights for digital marketing.

Saunders, Lewis and Thornhill (2019) divides thematic analysis into conventional and template analysis, which follows a more structured approach in which the research develops and adopts a codebook early on, revising it with each subsequent unit of data added in the process. Here, we follow the conventional approach of thematic analysis, which recommends the

researchers start coding before developing organizing structures (of codes, patterns, or themes). Braun and Clarke (2006, 2019) further developed the thematic analysis into a recursive and reflexive process that requires moving through data familiarization, coding, theme development, revision, naming, and writing up. Hence, this approach allows the researcher to identify concepts, makes generalizations, abstract ideas and uncovers patterns that could be developed into themes.

In fact, the search for themes is a common procedure in many qualitative data analysis, such as grounded theory, critical discourse analysis, qualitative content analysis, and narrative analysis (Bryman 2008). Hence, we discuss the qualitative analytical process before we introducing the six-phase process supporting the development of thematic analysis (Braun and Clarke 2013).

Qualitative analytical process

Thematic analysis aligns to a general qualitative research analytical process that comprises an ongoing and iterative interplay between the collection and analysis of data (Bryman 2008). Certainly, research programs offer specific procedures to navigate this process. Grounded theory and discourse analysis are two research programs with their own research traditions, methodological principles and strictly guided practice (Flick 2008). Nonetheless most research programs largely engage in three simultaneous activities: data condensation, data display and conclusion drawing/verification (Miles, Huberman and Saldaña 2018). Data condensation—the practice of selecting, focusing, simplifying, and/or transforming the data across the dataset (Miles, Huberman and Saldaña 2018)—holds together the ongoing and iterative analytical process which is pivotal to qualitative research inquiries. As noted,

“data condensation occurs continuously throughout the life of any qualitatively oriented

project. Even before the data are actually collected, anticipatory data condensation is occurring as the researcher decides (often without full awareness) which conceptual framework, which cases, which research questions, and which data collection approaches to choose. As data collection proceeds, further episodes of data condensation occur: writing summaries, coding, developing themes, generating categories, and writing analytic memos. The data condensing/transforming process continues after the fieldwork is over, until a final report is completed.” (Miles, Huberman and Saldaña 2018, 12).

Therefore the data condensation allows a more comprehensive understanding of the qualitative research analytical process than the widespread notion of data fragmentation/reorganization—which comprises of summarizing meaningful ideas to understand them, and thus repeat the work when wishing to undertake further analysis (Saunders, Lewis and Thornhill 2019). However, the fragmentation/reorganization process leaves unexplained what it takes to identify, work with, and then organize data into meaningful ideas.

The rationale around fragmentation/reorganization seems to be linked to the known analogy between quantitative research analysis and a puzzle assemblage (as seen in Le Compte, 2000; Saunders, Lewis and Thornhill 2019). In this analogy the researcher should be able to assemble the fragmented pieces (i.e. data excerpts) so s/he could see the whole picture. That is, create a comprehensive understanding of the social phenomena under investigation. Unlike the puzzle pieces, data excerpts do not come out ready from a box. Besides the decision about which chunk of data will be coded is already part of the analytical process. Fragmentation then becomes an important task in the coding process because it aids in the detangling of the raw data. More significantly, it allows the researcher to analyze the dataset, summarizing, distilling, or condensing the data, not simply reducing it (Saldaña 2013). Nonetheless, the researcher will contemplate layers of meanings in the data before selecting, focusing and simplifying it (Miles, Huberman and Saldaña, 2018). Hence, fragmentation is an essential task but does not denote the entire coding procedure.

Ultimately, a coding procedure is an active analytic process that requires an evolving and subjective interpretation of data (Braun and Clark 2019). The objectives in any coding procedure are twofold: ‘to develop and unfold an understanding of the issue or field under study first, which demands an open access to what should be coded and how; and secondly to identify an underlying structure, an organizing principle, a basic social process, or core category’ (Flick 2009, 436). Hence, the coding procedure involves two or more cycles. In the first cycle, coded data ranges from a single word to short passages or even a full page, whereas in the second cycle, coded data consists of similar segments, different/longer passages, analytic memos, besides a reorganization of the data coded thus far (Saldaña 2013). However, the mechanics of the coding procedure remains the same for this reason we are presenting it as an independent subsection.

The mechanics of coding procedure

The coding procedure involves scrutinizing collected data in order to relate it to theoretically relevant ideas, potential relationships, patterns of behavior, and, ultimately, research questions (Locke, Feldman and Golden-Biddle, 2020). It is then expected that several data segments (i.e. a word, short passages, image fragment) are identified and given a name that encapsulates their main idea—a code (Gibbs 2007, 2018). A code is “most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data” (Saldana 2013, 3). As such, codes work as labels that facilitate the retrieval of data that has been identified by the researcher as being similar in nature or meaning. Graham Gibbs (2018) offers one of the most didactical explanations about the coding procedure. Gibbs (2018) explains that coding, at its core, assists the researcher in reflecting on the data and its interpretation, but the actual coded

segment is only one aspect of that. Researchers new to coding often find it challenging to identify data segments and figure out what codes can represent these segments in a theoretical and analytic rather than merely descriptive manner. Those new to coding should carefully go through the data and decide what it is about before actually starting coding it (Gibbs 2018). One way to do so is through “intensive reading”: “In the visual arts the term ‘intensive seeing’ is used to refer to the way that we can pay close attention to all the things we can see, even the commonplace and ordinary. In the same way, you need to undertake ‘intensive reading’ when coding” (Gibbs 2018, 41). This activity is similar to the thematic analysis first phase, data familiarization (Braun and Clark 2012; 2019).

Gibbs’ (2018) developed his general qualitative research analytical procedure by successfully translating many techniques adopted by grounded theory researchers (Charmaz 2003, Glaser and Strauss 1967; Strauss and Corbin. 1997). We point out the sources used by Gibbs’ (2018) throughout this section. We draw from his explanation to introduce three mechanisms researchers can employ during the coding procedure: data interrogation, data sentence and data comparison.

Data interrogation. After the intensive reading, the researcher can start the coding procedure by employing the data interrogation mechanism. Its mechanic is rather simple, the researcher returns to the data and starts “asking” questions in order to recognize possible meanings and assign them a code. Based on Charmaz’s (2003, 94-5) work, Gibbs (2018) suggests a small list of basic questions a researcher can employ to interrogate the data:

- “What is going on?
- What are people doing?
- What is the person saying?
- What do these actions and statements take for granted?
- How do structure and context serve to support, maintain, impede or change these actions and statements?” (Charmaz 2003, 94-5)

At the beginning of the coding procedure, it is not uncommon for a researcher to assign more than one code to a data segment (Gibbs 2018). This is gradually resolved as the researcher starts searching for patterns in the coded data, clustering them together according to similarity and regularity (Saldaña 2013). Data interrogation then effectively help the researcher move away from descriptive to a more analytic and theoretical level of coding, especially when participants' own terms are coded (also known as in-vivo code). It is also useful in the development of themes. For instance, Saunders, Lewis and Thornhill (2019) create a set of questions for specific for this purpose (see examples in phase 3). Evidently, the researcher can develop his or her own set of question when seeking to interrogate the data.

Data sentience. The second mechanism for coding, data sentience, can aid in the analytic and theoretical interpretation of data. It does so by helping a researcher to develop sensitivity, knowledge and theoretical sophistication towards the data. To support researchers, Gibbs (2007, 2018) created a list of examples of what can be coded. He argues that while different authors have different priorities, many of the developed themes will be valuable in any coding procedure. More importantly, there is a common ground of phenomena that qualitative research programs will look for in their dataset. The researcher can resort to Gibbs' (2018) list when unable to grasp data meanings that are not obvious beforehand or codes that are simply descriptive and framed in terms of the respondents' worldviews.

Even though we borrow from Gibbs' (2018, 63) work the list of typical cases, we add our examples from Gonzalez-Arcos et al. (2021)'s study on consumer resistance to sustainability interventions.

What can be coded? *(with examples from the study on consumer resistance to sustainability interventions)*

1. **Specific acts, behaviors** – what people do or say.

Complaining to cashiers. Carrying bags to the trunk of their car.

2. **Events** – these are usually brief, one-off events or things someone has done. It is not uncommon for the respondent to tell them as a story.

Dropped all of their groceries at the checkout point. Started selling handmade reusable bags.

3. **Activities** – these are of longer duration than acts and often take place in a particular setting and may have several people involved.

Taking a sustainability workshop. Planning grocery shopping trips. Helping customers load their groceries onto reusable bags.

4. **Strategies, practices or tactics** – activities aimed towards some goal.

Keeping reusable bags in the car. Allocating a cabinet/drawer to bag storage. Purchasing biodegradable garbage bags.

5. **States** – general conditions experienced by people or found in organizations.

Anger, e.g. 'This is just another way for them to profit'

6. **Meanings** – a wide range of phenomena at the core of much qualitative analysis. Meanings and interpretations are important parts of what directs participants' actions.

- (a) What concepts do participants use to understand their world? What norms, values, rules and mores guide their actions?

Consumer choice as a framework to understand that disposable plastic bags should be made available, and that the government should be responsible for the environmental issue.

- (b) What meaning or significance does it have for participants, how do they construe events, what are their feelings?

Shame, e.g. 'I often forget to bring the bags and purchase new ones'

- (c) What symbols do people use to understand their situation? What names do they use for objects, events, persons, roles, settings and equipment?

Turtles or birds choking in plastic bag as a symbol of the environmental problem; "fabric bags" for reusable bags of any material.

7. **Participation** – people's involvement or adaptation to a setting.

Consumers getting involved in the intervention, e.g. bringing reusable bags to the grocery store

8. **Relationships or Interaction** – between people, considered simultaneously.

Collaboration in families or households to organize grocery shopping trips, e.g., 'my mom and my dad are the only ones doing the actual shopping'.

9. **Conditions or constraints** – the precursor to or cause of events or actions, things that reset behavior or actions.

Reduction in number of plastic bags offered for free in retail (max. 2); ban on disposable plastic bags.

10. **Consequences** – What happens If ...

Forgetting to bring reusable bags leads to additional expenses (e.g. need to purchase new bags) or shame.

11. **Settings** – the entire context of the events under study.

Supermarkets, other retailers. Households.

12. **Reflexive** – the researcher's role in the process, how intervention generated the data.

Expressing sympathy, e.g. 'It is a difficult change'; "liking" Facebook posts.

Gibbs's (2007, 2018) 'What can be coded?' list was adapted from Strauss (1987), Bogdan and Blieden (1992); and Mason (1996). As Gibbs notes, many of the points in the list are rather descriptive, but they were included because specific examples help people understand the phenomenon better.

Data comparison. As the coding procedure advances, the researcher becomes so familiar with ideas, relationships and behavioral patterns identified in the data that s/he misses what is significant (Gibbs 2007, 2018). Data comparison mechanisms bring out what is distinctive about the data and its content, helping to address this issue. Glaser and Strauss (1967) created the *constant comparison* mechanism, and it is still one of the cornerstones of the grounded theory approach. Nonetheless, coding with contrast or comparison techniques is used in many qualitative research analytical approaches (Saldana 2013). Building from Strauss, and

Corbin's (1990) work, Gibbs' (2018, 68-68) unfolds the data comparison mechanism into the following techniques²:

Analysis of word, phrase or sentence – Pick out one word or phrase that seems significant, then list all its possible meanings. Examine the text to see which apply here. You may find new meanings that were not obvious beforehand. Take, for example, the multiple meanings of “How do I carry all this now?”, stated in the title of Gonzalez-Arcos et al.'s (2021) article.

Flip-flop technique – Compare extremes on a dimension in question. For example, if someone mentions their income is a problem in shopping without disposable bags; contrast this with what it is like for those who have high income. You may discover dimensions or issues you hadn't thought of before.

Systematic comparison – Ask a series of ‘what ifs’ to explore all the dimensions of two phenomena. How would events, consequences, participation, etc., differ, in another setting or under different conditions? Gonzalez-Arcos and her coauthors (Gonzalez-Arcos et al. 2021) asked themselves, during the analytical process, ‘what if free biodegradable bags were offered to consumers?’; ‘what if disposable coffee cups were banned from cafes?’

Far-out comparisons – Take one element of the concept you are examining and think of the most remote or different example of some other phenomenon that shares some characteristics with that concept. Then work through all the other elements of both phenomena to see if they shed any light on the original. For example, you might compare consumer resistance to the ban on plastic bags to resistance to vaccination (i.e. Antivaxxer movement) to illuminate more aspects of the original idea.

Waving the red flag – Be sensitive to phrases like ‘Never’, ‘Always’, ‘It couldn't possibly be that way’. They are signals to look more closely. It is rarely the case that they are actually true. They usually mean things shouldn't happen that way. You need to find out what would happen if that situation actually did occur. (Gibbs 2018, 68-69)

These data comparison techniques are excellent ways to encourage more creative and in-depth thinking about the data, but as the researcher creates new codes, it is critical that s/he keep checking if previously coded data segments make sense now that s/he has done some more coding (Gibbs 2007; 2018).

Altogether, data interrogation, data sentience and data comparison strengthen the

² We adapted the techniques suggested by Gibbs (2018) to include examples from Gonzalez-Arcos et al.'s (2021) study.

iterative and evolving coding procedure and, ultimately, ensure analytical rigor. We advise readers to take a further look at Gibbs' (2007, 2018) work to gain a better understanding of the coding procedure.

It is important to have in mind that a coding procedure is essentially an interpretive act, not an exact science (Saldaña 2013). More significantly, the researcher need to align the coding procedure with the assumptions (i.e. ontology and epistemology) of their theoretical frameworks, for the generation of initial codes can be done in an inductive (i.e. emerging from the data, or bottom-up) or deductive (emergent from theory, or top-down) approach. Oftentimes coding is developed using a combination of these approaches, as researchers find emergent codes and adjust these to reflect existing concepts or start from theoretically inspired codes and adjust these to the dataset (abductive approach). Nonetheless, by carefully considering the assumption of their theoretical framework and how these connect with the phenomenon of interest, researchers should be able to justify their choice of analytical approach.

When working with large qualitative datasets, researchers may find useful to code with the assistance of qualitative data analysis software (e.g. Atlas ti, Nvivo³), but coding can also be done using comments and highlights on word processor software (e.g. Word, LaTeX) or by printing the dataset (i.e. interview transcripts, captions of digital data) and writing on the margins of these printouts. Whether coding manually or with assistance of computer software, the approach to thematic analysis is the same. Nonetheless, as Barbour (2009, 6) cautions, it is important “to learn the principles of qualitative data analysis before exploring any of the available computer packages. If you do not do this, then there is a very real danger that your

³ Lewins and Silver (2014) provide excellent advice on selecting the right computer-assisted qualitative data analysis software (QCDAs). Other useful information can be found at the following link:http://onlineqda.hud.ac.uk/which_software/what_packages_are_available/index.php

analyses are driven by the properties of the package rather than the other way around.” This is because data analysis software will come with tools that promise to facilitate theorizing—when, in fact, they simply alert the research to similarities identified based on the coding scheme developed by the researcher herself (Coffey, Beverley, and Paul 1996).

Overall, software-assisted coding will not necessarily be more rigorous than manual analysis. Most classic studies undertaking qualitative analysis were conducted without the use of computers. Furthermore, many researchers today do not have access to or cannot afford to use computer-assisted qualitative data analysis software (Gibbs 2007). The essential point to take away from all of this is that computer-assisted analysis isn't always necessary because the researcher, not the medium of analysis, is responsible for ensuring that the analysis is methodical and thorough (Barbour, 2009).

Thematic analysis in practice

To illustrate the development of thematic analysis, we continue to refer to the inquiry conducted by Gonzalez-Arcos and her coauthors (Gonzalez-Arcos et al. 2021) in examining how consumer practices change after sustainability interventions. We will work with this example to introduce the six steps of thematic analysis (Braun and Clarke 2012): (1) Developing familiarity with the data; (2) Generating initial codes; (3) Searching for themes; (4) Reviewing potential themes; (5) Defining and naming themes; (6) Producing a report.

Phase 1: Developing familiarity with the data

Gonzalez-Arcos and her coauthors (Gonzalez-Arcos et al. 2021) assembled a vast dataset of consumer reactions to the banning of plastic bags in Chile. Their dataset included tweets, Facebook posts and comments on those, news media articles, more than 50 interviews with

consumers and checkout assistants, and more than 60 observation reports created by consumers, including photographs and video clips. The researchers took on the task of familiarizing themselves with this dataset by holding meetings to look at the data together—each focusing on a specific source of data and sharing insights with the team. As most of the dataset was in Spanish, the non-Spanish speaking members of the research team would focus on the visual data, checking with the team when images alone were not sufficient to produce understanding. As they read through the dataset, the researchers were making notes, drawing connections between types of data, and exploring meaningful aspects of it that were connected to their research questions. One of the first aspects that caught the researchers’ attention was the emotional charge of comments about the sustainability intervention. Take, for example, this comment written by an individual in response to a supermarket post on Facebook reminding consumers to bring their own reusable bags, as disposable plastic bags would soon no longer be available: “The only thing they do is save money at the expense of customers, but their suppliers pack everything in plastic. Because customers must pay for looking after the environment...but this is hypocrisy because I have seen them throwing out plastic hangers and other things.”

The authors’ initial considerations about this and other similar pieces of evidence were that (a) consumers were angry about the ban, and (b) their anger was not directed at the government for passing the law, but at supermarkets, who were implementing the required measures. Gonzalez-Arcos and her coauthors (Gonzalez-Arcos et al. 2021) also noted, by immersing themselves in the dataset, that (c) consumers were making connections between the banning of plastic bags and “other things made of plastic” (Scaraboto n.d.), such as the plastic hangers mentioned in the quote above. In further reading the dataset, the researchers noted that not everyone was resistant: some people were plain confused about what would happen once the

plastic bag ban was implemented, and some others were enthusiastic about the ban. The authors also started to note that whereas interviewed consumers tended to be more conflicted or positive about the ban (“possibly due to social desirability bias from the interview process” [Scaraboto n.d.]), social media data and interviews with checkout assistants evidenced strong resistance against the ban. Along with each post or message about the law sent by the government, a supermarket, or an individual, came responses indicating discontent, anger, and objection.

Phase 2: Generating initial codes

Taking into consideration the initial observations developed in the first phase (and their research scenario, which includes a theoretical framework), researchers should generate initial codes that will facilitate the identification of patterns in the dataset. In our running example (Gonzalez-Arcos et al., 2021), the researchers started by developing a series of emergent codes that captured their initial observations about the dataset. By re-reading the dataset carefully, they developed a long list of codes which included: emotions (positive and negative); and meanings attributed to the plastic bag ban (e.g. good initiative, hypocrisy). At that stage, the researchers were already working with the possibility that the disruption caused by ban on plastic bags had made it very challenging for consumers to do their grocery shopping. Practice theories, which proposes that consumer behavior is not primarily determined by the individual but by the social practices through which they conduct their daily lives (e.g., eating, cooking, shopping), was adopted as the theoretical framework for the project. Gonzalez-Arcos and her colleagues (Gonzalez-Arcos et al. 2021) started to iterate between their early codes which emerged from the data and new codes based on the key concepts and relationships established in their practice theoretical framework: materials (e.g. plastic, bags, shopping cart); meanings (attached to the practice), and competences (e.g. actions and skills such as packing groceries, carrying bags), and

actors (those performing the practice). A summarized list of the codes generated through this abductive approach is presented on Table 1.

Code	Definition	Data Example
MEANINGS (OLDMNG)	Symbols, ideas, connections, aspirations related to the shopping practice prior to the ban	The polyethylene bags provide convenience during the transport of groceries, it is no less true that these have become the main pollution of Patagonia.
MEANINGS (NEWMNG)	Symbols, ideas, connections, aspirations related to the shopping practice after the ban	Good business for you Mr. businessman so that your profits prosper and the consumer do not profit. [Consumers] have to buy your bags and advertise your brand for free.
COMPETENCE (COMP)	Skill, know-how, and technique for shopping practice. Can be established, lost, or developing.	And how do you do it when you buy the groceries for the month? Does everything fit in the cart? And how do you do it to put the groceries in the cart when there are people waiting behind? Or do you not do that anymore? I ask out of curiosity because my family is big .I do not have a car so I have to think about how to do it with the cart.

Table 1 – Summary of codes and coded excerpts (Source: Scaraboto n.d.)

As coding, in thematic analysis, happens at an early stage of the analytical process, researchers should not be excessively cautious in coding only for aspects that are clearly connected to the research question. Interesting aspects of the dataset which, at this stage, may seem only tangentially related to the inquiry or to the theoretical framework adopted can also be coded for—and later on filtered out when themes are refined and defined.

Phase 3: Searching for themes

After having coded the entire dataset, researchers should start combining codes based on their similarity or fit, aiming to identify overarching themes that have the potential to provide meaningful insights about the phenomenon of interest. A theme is a category broader than codes (Saldaña 2013), and which indicates an important part of the answer to your research question, or an important insight about the phenomenon of interest.

The search for themes requires that researchers “brainstorm” around their long list of codes and make attempts to condense multiple codes into a coherent theme (though sometimes an important code can become a theme in itself). Saunders, Lewis and Thornhill (2019, p. 658) propose a series of questions that researchers should ask themselves as part of the analytical process of searching for themes in a dataset: “What are the key concepts in these codes? What, if anything, seems to be recurring in these codes? What seems to be important, whether it recurs often or not? What patterns and/or trends are evident in the coded data? Which codes appear to be related? How do a particular set of codes appear to be related?”.

In the example of the consumer resistance study, Gonzalez-Arcos and her colleagues (Gonzalez-Arcos et al. 2021) conducted several meetings to discuss the coded dataset. The researchers engaged in the search for themes by, first attempting to group existing codes together, and then, by identifying recurring patterns across codes. The authors found that, across codes there were manifestations of resistance, and that resistance seemed to be associated with the challenge of having to perform a changing social practice. In other words, consumers seemed to be angry about the ban of plastic bags because that required them to shop differently. The researchers identified several emergent themes connected to consumer resistance to sustainability interventions that required changes in social practices, including: “complicated competences”,

“contested meanings”, “competing materials”, and “unbundling” (See table 2 for this subset of potential themes and provisional definitions). Under “complicated competences”, for example, the researchers included codes related to frustration, anger, and competences, which captured consumers’ experience that shopping for groceries without free disposable plastic bags was inconvenient, difficult, and upsetting. Note how the themes adopt practice theory concepts (competences, meanings, materials) as a structure around which the researchers organize resistance manifestations. A figure was proposed (see figure 2) in an effort to organize these potential themes into a visual summary.

Subset of Potential Themes	Working Definitions
Complicated competences	Challenges in remembering to bring reusable bags, and other adjustments in shopping competence that were required from consumers.
Contested meanings	Meanings associated with the shopping practice became contested immediately after the ban.
Competing materials	Multiple substitute material replacements for the disposable bags that were distributed before.

Table 2 – Sample of potential themes (Source: Scaraboto n.d.)

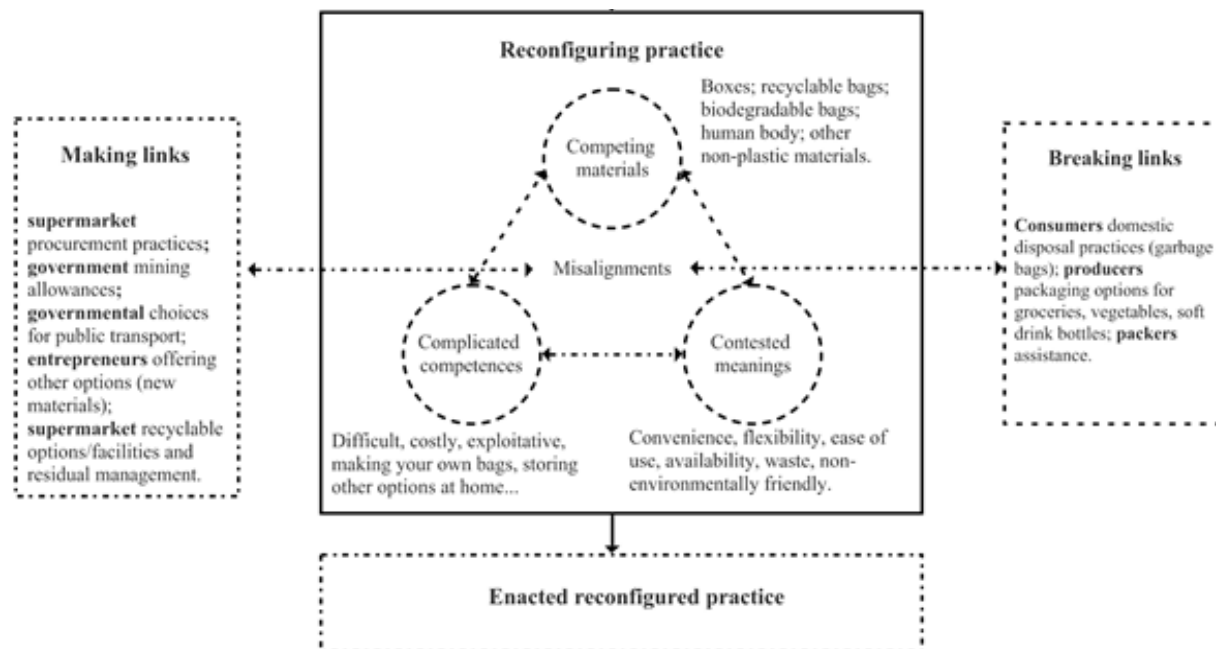


Figure 2 – Visual summary of potential research themes (Source: Scaraboto n.d.)

Phase 4: Reviewing potential themes

The next phase in the process of thematic analysis will have researchers reviewing the preliminary themes. This phase aims to produce a parsimonious and cohesive set of themes that are internally consistent, and externally distinctive. To achieve this, researchers must: (a) assess each of the themes for internal consistency by examining all of the coded data that is categorized under that theme; (b) verify that there is no overlap across themes, by considering whether coded excerpts could fit in multiple themes; and (c) assess that the group of themes is sufficient (and not excessive) to address the study’s research question(s). At this phase, researchers need to start visualizing themes together as a group of insights or answers that can explain the phenomenon of interest. It is very helpful, then, to elaborate thematic maps, diagrams, or figures that show how potential themes would fit together. Once such a visual summary is built, it becomes an “analytical framework” that should be tested against the dataset (or a subset of it) to verify whether the revised themes satisfactorily account for the dataset. Researchers can do so by

working with themes as if they were codes—and use them to code the data (it is important that this data is clear of previously employed codes). If the coding process unfolds without raising questions about where multiple excerpts should fit, and can accommodate a large part of the dataset, then the phase can be concluded. If not, another round of revision of potential themes may be needed.

In Gonzalez-Arcos et al.'s (2021) study, the authors found that, whereas initial themes were mostly internally consistent, there were overlaps across themes, suggesting that an underlying theme could be confounding the analysis. For example, the researchers noted that a form of conflicting emotionality was present across all themes. Anger, pride, frustration (codes that were previously under the theme “complicated competences”) also manifested in relation to competing materials, and contested meanings. Similarly, a discussion of responsabilization (i.e. questioning who was to “pay the price” for the elimination of disposable plastic bags) was evident in data excerpts that were coded and categorized under all three themes. As a consequence of this revision, the researchers decoupled “unsettling affectivity” (for emotion-related codes) and “responsibilization battles” (for blaming, disputing, and accusation codes) from the three themes they initially identified.

In assessing how this set of themes worked to address their research question, Gonzalez-Arcos et al. (2021) noticed that while themes structured at the level of practice elements (materials, competences and meanings) indicated changes to the practice that were brought about by the plastic bag ban, these themes did not describe what the ban required of consumers—and as such, did not explain why consumers resisted it so strongly. With this consideration in mind, the researchers revisited their theoretical framework and found that practice theories, in general, are largely silent about what individuals do when a practice is forced to change. As such, codes

derived directly from the theory would offer a limited explanation of the phenomenon. Gonzalez-Arcos et al (2021) then revisited their initial codes and reorganized them into three broad themes that consisted in categories of consumer actions: *making sense of the intervention*, *accommodating the change*, and *stabilizing the practice*. After the review, the researchers were convinced that the revised set of themes was coming together into an explanation for why consumers resisted sustainability interventions. The authors then checked this new set of themes against their dataset by coding additional data and assessing whether or not it fit the proposed framework. Figure 3 offers a visual summary of the revised themes.

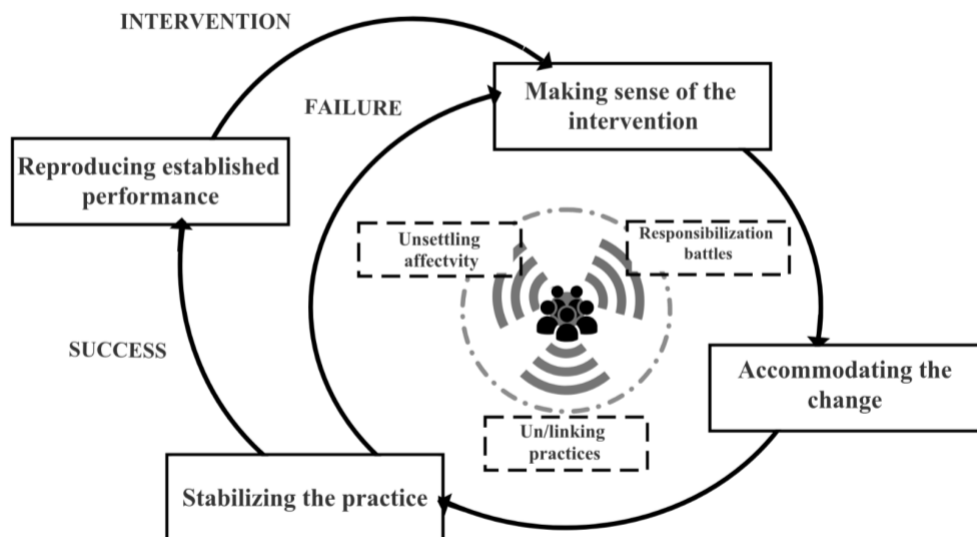


Figure 3 – Visual summary of revised themes as a framework (Source: Scaraboto n.d.)

Phase 5: Defining and naming themes

Once the research team has agreed on the thematic framework, it is important to clearly label and define each of the final themes. Labels should clearly communicate what the theme contains—either using a novel concept, or an existing concept derived from the project’s theoretical framework. In both cases, themes should be defined with precision to convey to readers what kind of data they could expect to find if looking at the excerpts categorized under

each theme. See, for example, the labels and definitions for the themes included in Gonzalez-Arcos and her colleagues' (Gonzalez-Arcos et al. 2021) thematic framework (Table 3).

Revised theme (Figure 3)	Label	Definition
Making sense of the intervention	Sensemaking	Consumers seek to understand and develop new meanings for the [reconfiguring] shopping practice.
Accommodating the change	Accommodating	Consumers develop new competences for using and handling the new materials (and meanings) involved in performing the shopping practice without disposable plastic bags
Stabilizing the practice	Stabilizing	Consumers embody (at times with resignation) the changed practice, with more or less difficulty or speed.
Responsibilization battles	Responsibilization battles	Consumers pushing back against being responsabilized for reconfiguring the shopping practice.
Unsettling affectivity	Unsettling emotionality	Consumers as carriers no longer feel completely attuned or “at home” with the shopping practice, which was previously familiar to them. The reconfiguration process becomes laden with unsettling emotions, including anxiety and fear.

Table 3 – Subset of revised themes with labels and definitions (Source: Scaraboto n.d.)

The researchers chose to simplify the labels for most of the themes, as shorter, action-focused labels more clearly communicated that consumers were intensively involved in the reconfiguration (i.e. transformation) of the shopping practice. Affectivity, a concept connected to practice theories, was replaced by “emotionality”, a term that can be more easily understood by readers who may not be well-versed in those theories, while still communicating the essence of the theme. Those labels that were considered satisfactory were kept, and definitions were added to each of the themes.

Phase 6: Producing a report

Producing a report for a study conducted using thematic analysis is an exercise of navigating between the complexity of the phenomenon and the (hopefully) simplicity of the thematic framework (or insights) generated as a result of the analysis. We recommend starting by selecting rich quotes—what Pratt (2008) calls *power quotes*: for each of the themes in the final thematic framework, select one or two compelling excerpts that best illustrate that theme. Consider creating one subsection of the report for each of the main themes, and using your “power quotes” in these subsections. Some researchers would like to start each sub-section could with the definition of the theme, followed by one of these rich quotes, which the researcher will unpack—that is, interpret and explain for the reader what is the connection between that quote, the theme, and how it partially contributes to answering the study’s research question. Other researchers may choose to start each subsection with a rich quote, and use it to explain how the theme is defined—and why it is labelled as such. Both approaches can be complemented by a selection of *proof quotes* (Pratt, 2008; 2009): more specific quotes that support each aspect of the argument you are making within a theme. These quotes can be presented in the body of text, if you do not have many themes to discuss. However, they can also be presented in a table, with as many “proof quotes” as needed to support your analysis of the theme. Some streamlining and organization of the themes can be adopted for communication purposes, as Gonzalez-Arcos et al. (2021, p.48) did: “When the ban challenged this shopping practice, we observed consumers seeking to change the practice through three reconfiguration processes. We present them separately for theorization but note that real-world reconfiguration processes are ongoing and recursive.”

Finally, in building an interpretation, it is very important to consider how each theme is

connected to the phenomenon of interest and the research questions, and add an explanation for how the themes come together to answer those questions. By developing those connections, researchers can move beyond findings that are “too descriptive”, a problem that often plagues research reports based on qualitative data, digital or not (Fischer, Gopaldas, and Scaraboto 2017). Theorizing grows from these explanations—as researchers attempt to make sense of their findings in light of the existing conceptual background and its limitations. For practitioners, insights can be developed in a similar fashion—by interpreting how current practice can be complemented or challenged by research findings. Although connections to existing theory and practice can be made in the findings sections as the themes are presented, an overall framework is often presented as the final subsection of the findings or in a separate Discussion section, and may consist in a theoretical extension.

In reporting findings from digital data, researchers often struggle to reproduce the richness of digital manifestations into the text format (still the most pervasive medium for research reports in academia and industry). When possible, we recommend trying to preserve visual elements that surround quoted excerpts (see Figure 4 for an example of how Gonzalez-Arcos et al. 2021 did so). An alternative is to create an appendix where quotes can be visualized as in their original format (e.g. as tweet or video screenshot).



Figure 4 – Images from dataset (Twitter posts). Source: (Gonzalez-Arcos et al. 2021)

Conclusion

Thematic analysis offers the researcher a flexible approach to engaging with digital qualitative data. As we note in this chapter, even though qualitative research in digital environments shares philosophies, data formats, and data analysis practices with traditional qualitative research, the research scenarios differ largely between these two types of inquiry. In fact, research scenarios may differ largely from one project to the next—particularly for those projects that are not restricted to one side of the increasingly crumbling boundaries of the online/offline domains. As such, one of the most important skills for qualitative researchers to develop when venturing in the collection and analysis of digital qualitative data is reflexivity (Alvesson, Sandberg, and Einola, forthcoming). Continued and consistent reflection about the phenomenon of interest, the researchers’ capacities and the theoretical frameworks available to them will likely keep researchers focused on collecting data that matters, and fine tune their analytical skills to identify themes that will lead to relevant insights (See Alvesson and

Sköldberg 2018 for principles that can be used as a basis for research reflexivity).

Overall, after having spent time with data in its raw form and worked through the phases of thematic analysis in an effort to build an interpretation of the phenomenon of interest, it is common that researchers feel that the thematic framework—and the theorization built based on it—does not capture all of the richness of digital data. At this point, it is important to remember that it is not possible to depict the complexity of social phenomena in all its nuance while abstracting enough to build relevant conceptual frameworks (Healy 2017). Well conducted thematic analysis can, nonetheless, offer researchers reassurance that they are fully leveraging digital data to generate insights and address specific research questions. Whether these insights and answers are relevant and timely, depends in large part on how a particular research scenario is set and evolves.

References

- Alvesson, M. & Sköldbberg, K. (2018), *Reflexive Methodology*. 3rd ed. London: Sage
- Alvesson, M., Sandberg, J., and Einola, K. (forthcoming), "Reflexive Design in Qualitative Research". In U. Flick (Ed), *The Sage Handbook of Qualitative Research Design*. London: Sage
- Barbour, R. (2009), "Analysis: Processing, Coding and Interrogating Data", *An Introduction to Qualitative Research*. Sage Publications, 1–20.
- Bogdan, R. and Biklen, S.K. (1997), *Qualitative research for education*, Allyn & Bacon.
- Braun, V. & Clarke, V. (2006), "Using thematic analysis in psychology", *Qualitative research in psychology*, 3(2) 77-101.
- Braun, V. & Clarke, V. (2019), "Reflecting on reflexive thematic analysis", *Qualitative Research in Sport, Exercise and Health*, 11(4) 589-597.
- Braun, V., & Clarke, V. (2012), "Thematic analysis". *APA handbook of research methods in psychology*, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological. American Psychological Association, 57–71
- Braun, V., & Clarke, V. (2013), *Successful qualitative research: A practical guide for beginners*, Sage Publication.
- Brüggen, E., & Willems, P. (2009), "A critical comparison of offline focus groups, online focus groups and e-Delphi." *International Journal of Market Research*, 51(3), 1-15.
- Catterall, M. and Maclaran, P. (2002), Researching consumers in virtual worlds: A cyberspace odyssey. *Journal of Consumer Behaviour: An International Research Review*, 1(3), 228-237.
- Charmaz, K. (2006) *Constructing grounded theory: A practical guide through qualitative analysis*, Sage Publication.
- Clarke, V., & Braun, V. (2014), "Thematic analysis", *Encyclopedia of critical psychology*, Springer, 1947-1952.
- Coffey, A., Beverley, H. and Paul, A., (1996), "Qualitative data analysis: Technologies and representations", *Sociological research online*, 1(1), 80-91.
- Denzin, N. K. & Lincoln, Y. S. (2008), "The Discipline and Practice of Qualitative Research," in *Collecting and interpreting qualitative materials*, Sage Publications, 1–44.
- Eagar, T. and Dann, S., (2016). "Capturing and analyzing social media composite content: the Instagram selfie", *Consumer culture theory*. Emerald Group Publishing Limited.
- Fischer, E., Gopaldas, A., & Scaraboto, D. (2017), "Why papers are rejected and how to get yours accepted", *Qualitative Market Research: An International Journal*, 10 (1), 60-67.
- Flick, U. (2008), *Designing Qualitative Research*, SAGE.
- Flick, U. (2009), *An Introduction to Qualitative Research*, Sage Publications (4th edition).
- Fox, N.J. and Alldred, P. (2015), "New materialist social inquiry: Designs, methods and the research-assemblage" *International Journal of Social Research Methodology*, 18(4) 99-414.
- Gaudet, S., and Robert, D. (2018), "Choosing Qualitative Inquiry," *A Journey Through Qualitative Research: From Design to Reporting*, Sage Publications, 1–19.
- Gibbs, G. R. (2018), "Thematic coding and categorizing," *Analyzing qualitative data* (Vol. 6). Sage Publications, 53-74.
- Gonzalez-Arcos-Arcos, C, Joubert, AM Scaraboto, Guesalaga, R., and Sandberg, J. (2021), "“How Do I Carry All This Now?” Understanding Consumer Resistance to Sustainability Interventions" *Journal of Marketing*, 85(3), 44-61.
- Healy, K.,(2017), Fuck nuance. *Sociological Theory*, 35(2), 118-127.

- Kozinets, R. V. (2019), *Netnography: The essential guide to qualitative social media research*, Sage Publications.
- Kvale, S., & Brinkmann, S. (2009), *Interviews: Learning the craft of qualitative research interviewing*, Sage Publications.
- Locke, K. D., Feldman, M., & Golden-Biddle, K. (2020), "Coding Practices and Iterativity: Beyond Templates for Analyzing Qualitative Data", *Organizational Research Methods*, 1-23.
- Mason, J. (1996), *Qualitative Researching*. Sage Publications.
- Paulus, T.M. and Lester, J.N. (2021). *Doing Qualitative Research in a Digital World*. Sage Publications.
- Plant, R. (2004), "Online communities." *Technology in Society*, 26(1), 51–65.
- 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*, 52 (5), 856-862.
- Reed, T. V. (2018), *Digitized lives: Culture, power and social change in the internet era*. Routledge.
- Saunders, Mark N, K., Lewis, P., and Thornhill, A. (2019), "Understanding research philosophy and approaches to theory development," in *Research Methods for Business Students*, Pearson Higher Ed.
- Silver, C. and Lewins, A., (2014). *Using software in qualitative research: A step-by-step guide*, Sage Publication.
- Scaraboto, D (n.d.), *Conversation among Daiane Scaraboto, Claudia Gonzalez-Arcos-Arcos, Alison M. Joubert, Rodrigo Guesalaga, and Jörgen Sandberg*.
- Schatzki, T. (2018), "On practice theory, or what's practices got to do (got to do) with it?", *Education in an Era of Schooling*, 151-165.
- Seale, C. (1999) "Quality in qualitative research." *Qualitative inquiry* 5 (4) 465-478.
- Shove, E., Pantzar, M. and Watson, M. (2012), *The dynamics of social practice: Everyday life and how it changes*. Sage.
- Silverman, D. (2013), *Doing Qualitative Research*.
- Strauss, A.L. (1987), *Qualitative analysis for social scientists*. Cambridge University Press.