

APPLYING THE CONTENT ANALYSIS METHOD IN CONSUMER BEHAVIOR RESEARCH

Asst. Prof. Alişan BALTACI

Yüksek İhtisas University, Vocational School of Health Services,
alisanbaltaci@yiu.edu.tr, ORCID: 0000-0002-3280-405X

Abstract

The usefulness of qualitative research methods attracts attention and interest in many fields today. For this reason, more and more researchers in the field of consumer behavior apply various qualitative research methods. One of the prominent qualitative research methods is "Content Analysis". In recent years, it has been seen that content analysis has been used in consumer behavior research, especially in examining data obtained through interviews or social media. Therefore, this research it is aimed to share basic information about how the content analysis method can be used in consumer behavior research in general.

Keywords: Qualitative Research, Content Analysis, Consumer Behavior

INTRODUCTION

Content analysis is generally the investigation of the content of a message within the framework of meaning, scope, and intent. The method is preferred by academics from various disciplines such as communication, psychology, political science, history, and language studies. From this point of view, it can be said that it is a method that is mostly used in the field of social sciences.

The method was first used more than 200 years ago to analyze textual content from newspaper and magazine articles, advertisements, political speeches, hymns, folktales, and riddles (Harwood & Garry, 2003).

Content analysis is about making valid, repeatable, and objective inferences about the message based on clear rules. Content analysis material can be letters, diaries, newspaper contents, ballads, short stories, radio, television messages, documents, texts, or symbols.

This research aims to draw the general conceptual framework for content analysis, conveying the basic information about how it can be applied. In addition, it is aimed to share the current situation in Turkish consumer behavior literature about content analysis.

LITERATURE REVIEW

1. Content Analysis

Content analysis has its roots in mass media research in the 1950s. These researches have generally been shaped within the framework of the inferences about the message in the sender, message, channel, and receiver processes in the basic communication model (White & Marsh, 2006). The prominent conceptual definitions in the literature regarding content analysis are presented in Table 1.

Table 1 A chronological overview of conceptual definitions of content analysis, Source: Bell et al., 2019; Prasad, 2008; Macnamara, 2005; Harwood and Garry, 2003

Researcher	Year	Definition
Berelson	1952	It is a research technique used to objectively, systematically, and quantitatively describe the explicit content of the communication.
Holsti	1968	It is a technique used to make inferences by systematically and objectively describing certain features of messages.
Krippendorff	1980	It is a research technique for making reproducible and valid inferences from data to their context.
Weber	1985	It is a research methodology that uses a set of procedures to draw valid inferences from the text.
Kerlinger	1986	It is a method of examining and analyzing communication in a systematic, objective, and quantitative way in order to measure variables.
Berger	1991	Content analysis is a research method used to evaluate the content (violence, negative depictions of women, etc.) of a famous art element in a representative sample.
Stone et al.	1996	Content analysis is an analysis method used to make inferences by systematically and objectively defining certain features in the text.
Neuman	1997	It is a technique used to gather and analyze text content. Content refers to words, meanings, images, symbols, ideas, themes, or any message that can be conveyed. On the other hand, the text describes anything written, visual, or spoken that serves as a means of communication.
Neuendorf	2002	Content analysis is a scientific qualitative research method that summarizes messages from measurable variables.
Collis & Hussey	2003	Content analysis is an analysis method that can be used with qualitative and quantitative data.
Bell et al.	2019	It is a qualitative approach that highlights the role of the researcher in the construction of texts and their meaning.

Although it is noteworthy that researchers in the literature generally refer to the definition made by Barelson in 1952, it can be said that the definition made by Neuman in 1997 and Neuendorf's definition in 2002 are the most comprehensive explanations in general. When the definitions are examined, it is understood that issues such as system, objectivity, quantity, context, and validity come to the fore.

Krippendorff (1980) suggests that six questions must be answered in every content analysis:

- 1) Which data are analyzed?
- 2) How are these data defined?

- 3) What is the population of the drawn data?
- 4) In which context are these data going to be analyzed?
- 5) What are the limits of analysis?
- 6) What is the goal of the inferences to be made? (Cited by: Stemler, 2000)

As can be understood from these six items, it is crucial to draw the research frame correctly before starting the content analysis studies.

The most crucial criticism of content analysis is about researcher bias and subjectivity. Researchers who make this criticism argue that an individual's interpretation of the meaning of particular content may lead to a situation that threatens objectivity. Such a situation may affect every stage of the research and its reliability (Harwood & Garry, 2003, Krippendorf, 2013). Carlson (2008) voices his criticism that even in content analysis studies that have been carried out adequately as a method, data can be misinterpreted, leading to wrong inferences. Another criticism of content analysis is that the method is too simple to explain complex phenomena (Morgan, 1993; Elo & Kyngas, 2008).

2. Content Analysis in Consumer Behavior Research

They must be understood by marketing professionals, as consumers participate in the value creation process (Rudd et al., 2018; Lecoivre et al., 2021). Therefore, businesses must listen to their customers to create shared experiences (Vargo & Lusch, 2017). At this point, it is critical to transforming the qualitative data obtained from consumers into information.

Another critical issue is the possibility of making faster and more accurate inferences using qualitative research software for consumer opinions, statements, and discourses, especially on social media. Today, consumers share a tremendous amount of accessible data via social media such as Twitter, Instagram, and Facebook (Gong et al., 2017). That provides an opportunity to access the consumers' experiences. Furthermore, each consumer's experience as a result of consumption may differ due to individual factors (Heinonen & Medberg, 2018). For this reason, content analysis can provide a broader perspective by defining consumers' individual experiences (Clarke et al., 1998).

Kassarjian's (1977) study, which is one of the most frequently referenced and one of the first studies in the literature on the use of content analysis in the field of consumer behavior, stated that the method could be used are as follows:

- What are the values that change periodically in society at the mass level?
- What are the images of the products and brands displayed in the mass media?
- What are the content features of the best-selling books? Can literary success be predicted based on these content characteristics?
- Which decision models are used by press and broadcast advertisers?

- Are minorities biased in fictional productions in the media?

Undoubtedly, in the current year 2022, considering the scope expansion in the field of consumer behavior since 1977, it would be appropriate to say that content analysis is now a tool that can be applied to more subjects. For example, considering that the symbolic value of products and brands is a factor that can affect the purchasing decision of consumers (Kırboğa, 2019). Therefore, it is considered that the content analysis method can also be used in research on symbolic consumption. Similarly, factors such as consumer attitudes and evaluations towards magazines, television, and internet advertisements can be examined through content analysis. In other words, content analysis can also be used as an effective method in the analysis of images or behaviors that contain implicit meanings.

Vespestad and Clancy (2021) examined the studies conducted with content analysis in the field of consumer behavior in their study between 1977 and 2017. Although the research has essential outputs, one of the prominent issues is that consumer behavior researchers often use the method to complement the quantitative method. However, it is not preferred to be used alone (Vespestad & Clancy, 2021).

2.1. Current Situation of Content Analysis in Turkish Consumer Behavior Literature

In order to determine the current state of Turkish literature, the first 100 pages of the Google Scholar search engine were investigated for ["content analysis" consumer]. As a result, 80 research articles conducted directly with consumer data were found. It is worth noting that the oldest study identified is from 1991, but the following oldest study is 12 years later. The distribution of the identified studies by year is shown in Table 2.

Table 2 Articles written using content analysis in the Turkish consumer behavior literature (1991-2022), Source: Google Scholar, 2022.

Year	Number of Research
1991	1
2013	1
2014	2
2015	2
2016	5
2017	11
2018	6
2019	13
2020	12
2021	25
2022	2
Total	80

Cite as: Baltacı, A. (2022). Tüketici Davranışları Araştırmalarında İçerik Analizi Yöntemi Kullanımı. Full Texts Book of International Congress on Social Sciences, China To Adriatic-XIV. M. Güneş Açıkgöz & Zhanuzak Alimgerey (Eds.). June 10-12, 2022, Kayseri. p. 153-167.

Although it is understood that the number of studies carried out in the field using consumer data has increased in general over the years, the year 2021 attracts particular attention. The data for 2022 is only two as of May. However, it is expected to increase by the end of the year.

The word cloud created by the MAXQDA qualitative data analysis program over the words in the titles of the 80 studies identified is given in Figure 1.



Figure 1 Word cloud for article titles written in the Turkish consumer behavior literature

As seen from the figure, research generally focuses on consumer comments or complaints on social media in the fields of travel and gastronomy. Another remarkable headline statement is COVID-19. That is because some of the studies conducted in 2020-21 were shaped around this subject. There are many different concepts, such as health, experience, and advertising, apart from those listed above. Another remarkable point is that the researchers mostly used the method for text-type data obtained from virtual media. The number of studies conducted using other qualitative data collection methods such as observation, interview, structured/semi-structured interview, visual content analysis, and brainstorming is almost non-existent. From this point of view, it can be said that there is a gap that can be filled for studies to be carried out with different subjects and data collection methods in future studies on Turkish literature.

3. Basic Principles in the Application of Content Analysis

Content analysis is used in clinical and social research to examine a wide range of content, from documents of interviews and discussions to the narratives and formats of films to TV

programs' editorial and advertising content of newspapers and magazines (Macnamara, 2005). Therefore, first, there is a need for content to be examined.

At this stage, it is considered that it would be appropriate to define three concepts that are frequently encountered in content analysis research. These concepts are shared in Table 3.

Table 3 Code, category, and theme, Source: Erlingsson & Brysiewicz, 2017.

Concept	Definition
Code	Code can be thought of as a tag. It is a dense definition that is usually one or two words long and expresses the meaning of a unit.
Category	Categories emerge by grouping codes related to each other in terms of content or context. A category generally answers the questions of who, what, when, or where.
Theme	A theme answers questions such as why, how, in what way, or in what sense. The theme expresses a basic meaning found in two or more categories.

The content analysis begins with the researcher clarifying questions about the content at hand. For this, the researcher asks, "What do I want to learn from this content?" and then shapes the research within the framework of the answer to this question (Prasad, 2008).

Although the different numbers and titles of process steps regarding content analysis have been defined in the literature, it can be said that the method generally consists of 8 process steps. These process steps can be seen in Figure 2.

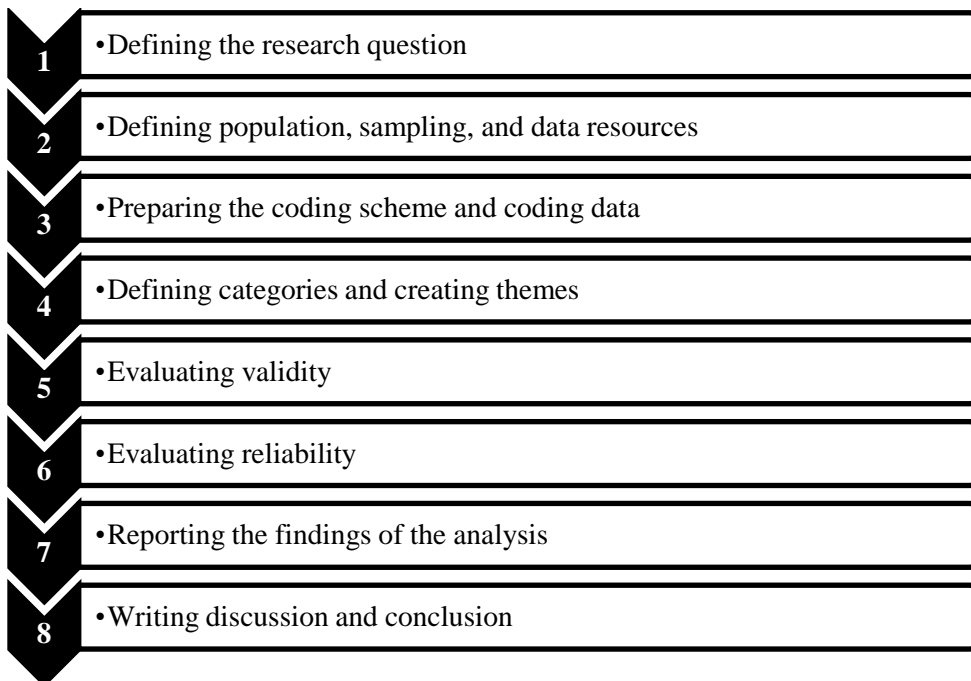


Figure 2 Process steps in content analysis

Content analysis can be performed inductively and deductively, depending on the need. The inductive method represents a discovery process. The inductive approach is mostly needed when the literature on the phenomenon is insufficient, or the content is fragmented, and an

exploratory approach is required (Elo & Kyngas, 2008). On the other hand, the deductive method is mostly used to test a specific theory in different situations or times (Krippendorf, 2013).

Content analysis can generally be carried out in three approaches:

- Human-rated: In this approach, the researcher manually encodes according to the developed coding scheme.
- Individual word counting systems: In this approach, the researcher creates categories by bringing together synonyms to reveal a particular interest in specific content. At this stage, relative importance levels of words emerge according to the frequency level.
- Computerized systems using artificial intelligence: In this approach, the researcher uses computers and software to analyze the content within the framework of specific algorithms (Short & Palmer, 2008). Nowadays, researchers can carry out their research with various computer-aided analysis software such as NVivo, QDA, WordStat, etc.

Since content analysis is a very flexible method, researchers can perform their analyses by choosing at least one of the above approaches or different approaches suitable for their needs.

3.1. Defining the Research Question

The research question is the heart of the research. It is a clear statement about the problem that the researcher wants to contribute to the solution through the relevant research (Nenty, 2009). At this stage, the researcher needs to seek answers to the following questions.

- What is the problem I am trying to contribute to the solution? What are the borders, causes, history of the problem, and who are its stakeholders?
- Why is it important to solve this problem? Why is the issue I am trying to contribute to a solution a problem? Who is affected by this problem? What could be the negative consequences if the problem is not resolved? Are consumers the only ones affected by these results?
- What kind of contribution can I make to the theory and practice in the field of consumer behavior as a result of this research? How can other problems in the field of consumer behavior be solved by practitioners and theorists by using my solution?

When the research question is considered in this context, it also determines the content and limits of the purpose of the research.

After clarifying the research question in the quantitative analysis method, it usually starts with writing the hypotheses. A hypothesis is a tentative proposition that explains a set of facts and can be tested in further research. In other words, it is a proposition that expresses the expected

relationship between the variables that are the subject of the research (Ingham-Broomfield, 2014). Hypotheses must be testable. Content analysis is a qualitative research method and is a process related to examining theoretical expectations rather than testing hypotheses. In qualitative research, data is not for testing a hypothesis; they are interpreted and analyzed to ensure results are reached (Draper, 2004). Therefore, there is no hypothesis test in content analysis due to its qualitative nature, but this does not eliminate the necessity of the research having a theoretical purpose (Haven & Van Grootel, 2019). For this reason, researchers should determine their theoretical expectations from the research at this stage.

3.2. Defining Population, Sampling, and Data Resources

In qualitative and quantitative studies, sample selection is made from a pool that can represent the population and is suitable for the research. The difference in qualitative research is that the selected sample can shift in different directions or increase in-depth as it is examined. On the other hand, qualitative research aims to transfer rather than generalize. For this reason, the probability of being included in the sampling or estimation of each unit in the population does not have to be equal. Furthermore, since sample selection and coding process can be handled simultaneously, there is also the possibility and opportunity to make changes in sampling selection during coding (White & Marsh, 2006).

Although there are many different sampling methods, it is said that three sampling methods stand out in general in qualitative studies (Marshall, 1996).

Table 4 The most frequently used sampling methods in content analysis, Marshall, 1996; Yagar and Bulk, 2018.

Sampling Method	Definition
Convenience Sampling	It is a method that aims to reach the most accessible data. It is the cheapest in terms of time and financial costs. However, there is a risk of producing low-quality data because of its poor reliability.
Judgement Sampling	It is the most common method. The researcher chooses the most efficient sample for the research question in this method. Although demographic variables are essential, participants can also be included in the sample according to the socio-economic and psychological aspects of their expertise or experience. In some cases, the number of participants can be increased by applying the snowball method.
Theoretical Sampling	This method enables the discovery of grounded theory by collecting empirical data and comparing them repeatedly in the light of theoretical parameters until they no longer produce a different meaning.

Finally, it should be noted that the most crucial feature of a good sample is that it can answer the question the researcher seeks to answer (Kassarjian, 1977). Therefore, researchers should determine a method suitable for their research design, test the data they obtained from the sample, and update their methods if they detect nonconformity in the tested data.

The definition of the population can be made through individuals among consumer groups, or it can be selected through many different channels such as advertising content, customer complaints, salesperson opinions, and newspaper news. Therefore, the sample will ultimately be defined over the selected population. In parallel, data sources can be obtained through sources such as people, various texts, images, or sounds. The critical issue is that the selected data source contains the answer to the research question.

3.3. Preparing the Coding Scheme and Coding Data

The coding scheme is created based on reviewing the relevant literature and research questions (Harwood & Garry, 2003). A coding scheme is a classification tool that the researcher uses to code the expressions in the data set. This tool should be simple and articulate. Thus, the probability of people who will code similar concepts in the future may apply the same expressions, and that would increase the schema's reliability (White & Marsh, 2006).

The creation and implementation of the coding scheme require an eight-step process. This process is shown in Table 5.

Table 5 Creation and implementation of the coding scheme, Source: Weber, 1990.

Step	Action
Defining coding units	The units to be coded, such as word, sentence, paragraph, etc., should be defined.
Defining categories	Categories should not be defined too broadly. For example, "Consumer Expectation" is an extensive category and includes many elements. Whereas conceptually delimited category definitions such as "Consumer Price Expectation", "Consumer Quality Expectation", "Consumer Performance Expectation", etc., would be more appropriate.
Coding on sample content	A small portion of the content is coded and tested so that the categories can be clearly defined. This action reveals possible uncertainties and makes it possible to update the prepared coding scheme and make it healthier for the following research step.
Evaluation of accuracy and reliability	That is the stage for confirming the coded content's accuracy and the coding's reliability.
Review of coding rules	The coding rules should be reviewed and re-arranged if the reliability is low or errors are detected.
Back to coding on the sample text	The process should be repeated until the accuracy and reliability are reasonable.
Encoding the entire content	The coding scheme and rules can be applied to the entire content when a reasonable level of accuracy and reliability is achieved.
Evaluation of ultimate accuracy and reliability	Although a general evaluation has been made of the sample content, it should be confirmed that all content is accurate and reliable at the expected level.

In content analysis, coding schemes created in previous studies can be applied, or researchers can develop their own schemes according to needs (Krippendorff, 2013). In addition, changing the coding schemes applied in previous research can help explain the latent contents

(Harwood & Garry, 2003). The codes briefly describe the condensed meaning units and help the researcher to form a new perspective on the data. Researchers can make as many changes as they want until they are sure that the codes are correct (Erlingsson & Brysiewicz, 2017). It may be helpful in the later stages of the research for the researcher to take notes about his findings during the coding phase. On the other hand, it should be kept in mind that it may be beneficial to consult experts with in-depth knowledge in the field while creating coding schemes, especially in exploratory research.

Feng et al.'s (2019) coding scheme is shown as an example in Table 6. That research examines consumer opinions about a campaign on YouTube. Researchers used java-based software (YouTube Comment Scraper) to obtain and analyze the data.

Table 6 Sample coding scheme, Source: Feng et al., 2019

Code	Definition
Ad skepticism	The consumer criticizes the advertising video for various reasons in his comments.
Beauty definition	The consumer discusses the definition of beauty from different perspectives in his interpretation.
Praise	The consumer makes positive evaluations of the advertising video in his comments.
Discussion of broad issues	The consumer not only discusses the ad video in his comment but also discusses some social and cultural problems such as stereotypes, gender, race, etc.
Other	The consumer discusses insignificant details such as the background music of the advertising video or other irrelevant issues in his comments.

3.4. Defining Categories and Creating Themes

After the coding process, the codes are divided into categories to answer the questions of who, where, when, and what. In other words, categories are created by combining codes that answer similar questions (Erlingsson & Brysiewicz, 2017).

Themes are created by combining two or more categories to reveal when the dataset contains hidden meanings. Themes are created to answer questions such as why, how, in what way, and in what sense (Erlingsson & Brysiewicz, 2017).

3.5. Evaluating Validity

In quantitative content analysis, the coding scheme is a list of coding rules determined before the research begins. Validity is a concept related to the extent to which a measurement rule represents the intended phenomenon (Neuendorf, 2005). In the literature, it is seen that several different methods related to validity are used in content analysis:

- **Face Validity:** It is a type of validity related to the level of appearing to measure what a measurement tool claims to measure (Ayyıldız & Tarhan, 2014). In the literature, it is seen that expert opinion is frequently sought in order to ensure face validity.
- **Criterion Validity:** It is a comparative method that looks at the compatibility of the measurement tool with similar measurement tools (Çalışgan, 2015). It is based on the assessment of conformity between the code and the criteria.
- **Construct Validity:** It is a type of validity that examines the extent to which a measure is related to other measures (constructs) in a way that is consistent with the constructs produced from theory (Neuendorf, 2005). It is much more difficult to detect than content or criterion validity, but it is imperative.

3.6. Evaluating Reliability

Krippendorff (1980) mentions three different types of reliability in content analysis. These are shown in Table 7 (Cited by Harwood & Garry, 2003).

Table 7 Reliability in content analysis, Source: Krippendorff, 1980; Cited by Harwood and Garry, 2003

Type	Definition
Consistency	It means that the same coder does not make a different judgment when coding the same data at different times. This situation can also be defined as the invariance of analysis. However, it is the weakest form of reliability.
Repeatability	Repeatability means there should be a consensus even when different people under different conditions do coding.
Accuracy	It refers to the process of conforming to known standards. The errors that occur in this case are due to coding inconsistencies.

Based on the table, it can be said that researchers can increase reliability in the content analysis process if they comply with the generally accepted research processes, have knowledge of the literature, and determine the codes and categories with sensitivity clearly and understandably.

Based on the table, the answers to the following questions in a content analysis can give an idea about the reliability level of the research:

- **Consistency or single coder reliability:** Does the same coder get the same results every time they encode the same content?
- **Repeatability or cross-coder reliability:** Do coding schemes ensure that different people code the same content within the same categories?

If the researcher answers yes to both questions, it will indicate an acceptable situation regarding reliability (Stemler, 2000).

In addition, the researchers can also benefit from numerical measurement and evaluation tools such as Cohen's κ (kappa) coefficient, Scott's π coefficient, Krippendorff's α coefficient, Bennett, Alpert, and Goldstein's S coefficient, etc., to test reliability (Oleinik et al., 2014).

Duriau et al. (2007) and Lombard et al. (2002) examined the articles that applied content analysis. Lombard et al. (2002) found 31%, and Duriau et al. (2007) found that 38% of the researchers did not clearly state the reliability issues in their research. Therefore, it would be appropriate to say that the issue of reliability is an essential but neglected component of the content analysis.

3.7. Reporting the Findings of the Analysis

After coding, creating themes and categories, and determining reliability and validity, inferences about the findings will need to be reported. Content analysts use tables, crosstabs, graphs, clustering, relationship analysis, etc. can benefit from many statistical and non-statistical analysis methods (White & Marsh, 2006). The data quality and the research's purpose are vital in choosing the method to apply. Reporting can be done with one or more methods to provide the most appropriate output for the research from the data at hand. At this point, research design will play an essential role in preferences.

On the other hand, using the contents of the data examined in the transfer of the obtained findings will contribute to understanding the situation (Yin, 2016). An example of this transfer method is shared below:

Bateson and Hui (1987) state that increased perceived crowding will create dissatisfaction with the shopping environment. That may cause the individual to avoid shopping or leaving the shopping environment. Regarding this issue, the participant (28 years old, Female) says, *“I cannot shop in crowded environments. Sometimes I think everyone is looking at me and condemning me. This makes me feel bad”* (Baltacı, 2019).

This way, the link between the literature and the data examined and how the coding scheme is used can be conveyed to the reader more effectively.

3.8. Writing Discussion and Conclusion

Before concluding the research, the researcher makes objective and critical comments on the analysis he has referred to in the discussion section. Finally, in conclusion, all the significant findings related to the primary purpose of the research and the level of answering the research question should be mentioned (Ingham-Broomfield, 2014).

The topics covered in the results section of the research should improve the theoretical framework or contribute to developing a new theory/model (Johnson et al., 2020). It will not be surprising that studies that go beyond the collected data and can affect different studies with their findings and results are ahead of the others (Yin, 2016). For this reason, one of the critical sections that shape the future of research is the discussion and results section.

Another critical issue is that this section should contain information for future users, such as how practitioners and researchers can use the information obtained from the research. Also, what other information can be produced by using the results and findings of the research should be mentioned. In other words, in the discussion and conclusion part, the general picture of the research should be drawn, and a short road map for the future should be given after the significant findings are discussed.

DISCUSSION AND CONCLUSION

Content analysis is an important method that can be applied to theoretically explain individuals' consumption behavior in the consumer behavior field, where the subject is human. In addition, the increased number and quality of software for qualitative data analysis in recent years also provides a significant opportunity for researchers.

Today, on social media, consumers are interested in products, brands, statuses, advertisements, etc. It is known that he freely shares his thoughts on many marketing elements and communicates with other consumers. Of course, digitizing such a large amount of data is not easy. However, it is quite possible to subject this accessible data to content analysis. Undoubtedly, social media is not the only point where consumer behavior researchers can access data. It is possible to reach qualitative data with many different methods such as interviews, observation, brainstorming, etc.

At this point, the important thing is to draw the research boundaries clearly and create the coding scheme that can produce the most meaningful data within these limits. Researchers should avoid subjectivity and act impartially in creating the coding scheme and analyzing the data. Particular attention should be paid to validity and reliability. Otherwise, there is a possibility of producing erroneous or meaningless results by falling into research bias.

Finally, with this study, the conceptual framework of content analysis, which is used as an essential information generation tool in the field of consumer behavior, is given, and the processes that practitioners should apply are examined. Due to the dynamic nature of the content analysis method and the field of consumer behavior, it would be appropriate to think that the scope of the method and the richness of its application will develop over time. At this point, the responsibility of the researchers is to follow the current literature and stay up-to-date on new application possibilities.

REFERENCES

[Click here to see the original conference paper for references.](#)