


Qualitative Content Analysis: A Focus on Trustworthiness

SAGE Open
January-March 2014: 1–10
© The Author(s) 2014
DOI: 10.1177/2158244014522633
sgo.sagepub.com


Satu Elo¹, Maria Kääriäinen^{1,2}, Outi Kanste³, Tarja Pölkki¹, Kati Utriainen¹, and Helvi Kyngäs^{1,2}

Abstract

Qualitative content analysis is commonly used for analyzing qualitative data. However, few articles have examined the trustworthiness of its use in nursing science studies. The trustworthiness of qualitative content analysis is often presented by using terms such as *credibility*, *dependability*, *conformability*, *transferability*, and *authenticity*. This article focuses on trustworthiness based on a review of previous studies, our own experiences, and methodological textbooks. Trustworthiness was described for the main qualitative content analysis phases from data collection to reporting of the results. We concluded that it is important to scrutinize the trustworthiness of every phase of the analysis process, including the preparation, organization, and reporting of results. Together, these phases should give a reader a clear indication of the overall trustworthiness of the study. Based on our findings, we compiled a checklist for researchers attempting to improve the trustworthiness of a content analysis study. The discussion in this article helps to clarify how content analysis should be reported in a valid and understandable manner, which would be of particular benefit to reviewers of scientific articles. Furthermore, we discuss that it is often difficult to evaluate the trustworthiness of qualitative content analysis studies because of defective data collection method description and/or analysis description.

Keywords

analysis method, nursing methodology research, qualitative content analysis, qualitative research, rigor, trustworthiness, validity

Although qualitative content analysis is commonly used in nursing science research, the trustworthiness of its use has not yet been systematically evaluated. There is an ongoing demand for effective and straightforward strategies for evaluating content analysis studies. A more focused discussion about the quality of qualitative content analysis findings is also needed, particularly as several articles have been published on the validity and reliability of quantitative content analysis (Neuendorf, 2011; Potter & Levine-Donnerstein, 1999; Rourke & Anderson, 2004) than qualitative content analysis. Whereas many standardized procedures are available for performing quantitative content analysis (Baxter, 2009), this is not the case for qualitative content analysis.

Qualitative content analysis is one of the several qualitative methods currently available for analyzing data and interpreting its meaning (Schreier, 2012). As a research method, it represents a systematic and objective means of describing and quantifying phenomena (Downe-Wamboldt, 1992; Schreier, 2012). A prerequisite for successful content analysis is that data can be reduced to concepts that describe the research phenomenon (Cavanagh, 1997; Elo & Kyngäs, 2008; Hsieh & Shannon, 2005) by creating categories, concepts, a model, conceptual system, or conceptual map (Elo & Kyngäs, 2008; Morgan, 1993; Weber, 1990). The research

question specifies what to analyze and what to create (Elo & Kyngäs, 2008; Schreier, 2012). In qualitative content analysis, the abstraction process is the stage during which concepts are created. Usually, some aspects of the process can be readily described, but it also partially depends on the researcher's insight or intuitive action, which may be very difficult to describe to others (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004). From the perspective of validity, it is important to report how the results were created. Readers should be able to clearly follow the analysis and resulting conclusions (Schreier, 2012).

Qualitative content analysis can be used in either an inductive or a deductive way. Both inductive and deductive content analysis processes involve three main phases: preparation, organization, and reporting of results. The preparation

¹University of Oulu, Finland

²Medical Research Center, Oulu University Hospital, Finland

³National Institute of Health and Welfare, Oulu, Finland

Corresponding Author:

Satu Elo, Senior University Lecturer, Institute of Health Sciences, Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Box 5000, 90014, Finland.
Email: satu.elo@oulu.fi

phase consists of collecting suitable data for content analysis, making sense of the data, and selecting the unit of analysis. In the inductive approach, the organization phase includes open coding, creating categories, and abstraction (Elo & Kyngäs, 2008). In deductive content analysis, the organization phase involves categorization matrix development, whereby all the data are reviewed for content and coded for correspondence to or exemplification of the identified categories (Polit & Beck, 2012). The categorization matrix can be regarded as valid if the categories adequately represent the concepts, and from the viewpoint of validity, the categorization matrix accurately captures what was intended (Schreier, 2012). In the reporting phase, results are described by the content of the categories describing the phenomenon using a selected approach (either deductive or inductive).

There has been much debate about the most appropriate terms (rigor, validity, reliability, trustworthiness) for assessing qualitative research validity (Koch & Harrington, 1998). Criteria for reliability and validity are used in both quantitative and qualitative studies when assessing the credibility (Emden & Sandelowski, 1999; Koch & Harrington, 1998; Ryan-Nicholls & Will, 2009). Such terms are mainly rooted in a positivist conception of research. According to Schreier (2012), there is no clear dividing line between qualitative and quantitative content analysis, and similar terms and criteria for reliability and validity are often used. Researchers have mainly used qualitative criteria when evaluating aspects of validity in content analysis (Kyngäs et al., 2011). The most widely used criteria for evaluating qualitative content analysis are those developed by Lincoln and Guba (1985). They used the term *trustworthiness*. The aim of trustworthiness in a qualitative inquiry is to support the argument that the inquiry's findings are "worth paying attention to" (Lincoln & Guba, 1985). This is especially important when using inductive content analysis as categories are created from the raw data without a theory-based categorization matrix. Thus, we decided to use such traditional qualitative research terms when identifying factors affecting the trustworthiness of data collection, analysis, and presentation of the results of content analysis.

Several other trustworthiness evaluation criteria have been proposed for qualitative studies (Emden, Hancock, Schubert, & Darbyshire, 2001; Lincoln & Guba, 1985; Neuendorf, 2002; Polit & Beck, 2012; Schreier, 2012). However, a common feature of these criteria is that they aspire to support the trustworthiness by reporting the process of content analysis accurately. Lincoln and Guba (1985) have proposed four alternatives for assessing the trustworthiness of qualitative research, that is, credibility, dependability, conformability, and transferability. In 1994, the authors added a fifth criterion referred to as authenticity. From the perspective of establishing credibility, researchers must ensure that those participating in research are

identified and described accurately. Dependability refers to the stability of data over time and under different conditions. Conformability refers to the objectivity, that is, the potential for congruence between two or more independent people about the data's accuracy, relevance, or meaning. Transferability refers to the potential for extrapolation. It relies on the reasoning that findings can be generalized or transferred to other settings or groups. The last criterion, authenticity, refers to the extent to which researchers, fairly and faithfully, show a range of realities (Lincoln & Guba, 1985; Polit & Beck, 2012)

Researchers often struggle with problems that compromise the trustworthiness of qualitative research findings (de Casterlé, Gastmans, Bryon, & Denier, 2012). The aim of the study described in this article was to describe trustworthiness based on the main qualitative content analysis phases, and to compile a checklist for evaluating trustworthiness of content analysis study. The primary research question was, "What is essential for researchers attempting to improve the trustworthiness of a content analysis study in each phase?" The knowledge presented was identified from a narrative literature review of earlier studies, our own experiences, and methodological textbooks. A combined search of Medline (Ovid) and CINAHL (EBSCO) was conducted, using the following key words: trustworthiness, rigor OR validity, AND qualitative content analysis. The following were used as inclusion criteria: methodological articles focused on qualitative content analysis in the area of health sciences published in English and with no restrictions on year. The search identified 12 methodological content analysis articles from databases and reference list checks (Cavanagh, 1997; Downe-Wamboldt, 1992; Elo & Kyngäs, 2008; Graneheim & Lundman, 2004; Guthrie, Yongvanich, & Ricceri, 2004; Harwood & Garry, 2003; Holdford, 2008; Hsieh & Shannon, 2005; Morgan, 1993; Potter & Levine-Donnerstein, 1999; Rourke & Anderson, 2004; Vaismoradi, Bondas, & Turunen, 2013). The reference list of selected papers was also checked, and qualitative research methodology textbooks were used when writing the synthesis of the review. The discussion in this article helps to clarify how content analysis should be reported in a valid and understandable manner, which, we expect, will be of particular benefit to reviewers of scientific articles.

Trustworthiness in the Preparation Phase in Content Analysis Study

Based on the results of the literature search, the main trustworthiness issues in the preparation phases were identified as trustworthiness of the data collection method, sampling strategy, and the selection of a suitable unit of analysis. Based on the findings, we have compiled a checklist for researchers attempting to improve the trustworthiness of a content analysis study in each phase (Table 1).

Table 1. Checklist for Researchers Attempting to Improve the Trustworthiness of a Content Analysis Study.

Phase of the content analysis study	Questions to check
Preparation phase	Data collection method
	How do I collect the most suitable data for my content analysis?
	Is this method the best available to answer the target research question?
	Should I use either descriptive or semi-structured questions?
	Self-awareness: what are my skills as a researcher?
	How do I pre-test my data collection method?
	Sampling strategy
	What is the best sampling method for my study?
	Who are the best informants for my study?
	What criteria should be used to select the participants?
Organization phase	Is my sample appropriate?
	Is my data well saturated?
	Selecting the unit of analysis
	What is the unit of analysis?
	Is the unit of analysis too narrow or too broad?
	Categorization and abstraction
	How should the concepts or categories be created?
	Is there still too many concepts?
	Is there any overlap between categories?
	Interpretation
What is the degree of interpretation in the analysis?	
How do I ensure that the data accurately represent the information that the participants provided?	
Reporting phase	Representativeness
	How to I check the trustworthiness of the analysis process?
	How do I check the representativeness of the data as a whole?
	Reporting results
	Are the results reported systematically and logically?
	How are connections between the data and results reported?
	Is the content and structure of concepts presented in a clear and understandable way?
	Can the reader evaluate the transferability of the results (are the data, sampling method, and participants described in a detailed manner)?
	Are quotations used systematically?
	How well do the categories cover the data?
Are there similarities within and differences between categories?	
Is scientific language used to convey the results?	
Reporting analysis process	
Is there a full description of the analysis process?	
Is the trustworthiness of the content analysis discussed based on some criteria?	

Data Collection Method

Demonstration of the trustworthiness of data collection is one aspect that supports a researcher's ultimate argument concerning the trustworthiness of a study (Rourke & Anderson, 2004). Selection of the most appropriate method of data collection is essential for ensuring the credibility of content analysis (Graneheim & Lundman, 2004). Credibility deals with the focus of the research and refers to the confidence in how well the data address the intended focus (Polit & Beck, 2012). Thus, the researcher should put a lot of thought into how to collect the most suitable data for content analysis. The strategy to ensure trustworthiness of content analysis starts by choosing the best data collection method to answer the research questions of interest. In most studies

where content analysis is used, the collected data are unstructured (Elo & Kyngäs, 2008; Neuendorf, 2002; Sandelowski, 1995b), gathered by methods such as interviews, observations, diaries, other written documents, or a combination of different methods. However, depending on the aim of the study, the collected data may be open and semi-structured. If inductive content analysis is used, it is important that the data are as unstructured as possible (Dey, 1993; Neuendorf, 2002).

From the perspective of trustworthiness, a key question is, "What is the relationship between prefiguration and the data collection method, that is, should the researcher use descriptive or semi-structured questions?" Nowadays, qualitative content analysis is most often applied to verbal data such as interview transcripts (Schreier, 2012). With descriptive data

collection, it can often be challenging to control the diversity of experiences and prevent interviewer bias and the privileging of one type of information or analytical perspective (Warr & Pyett, 1999). For example, when using a descriptive question such as “Could you please tell me, how do you take care of yourself?” the researcher has to consider the aim of data collection and try to extract data for that purpose. However, if the researcher opts for a semi-structured data collection method, they should be careful not to steer the participant’s answers too much to obtain inductive data. It may be useful for the interview questions to be developed in association with a “critical reference group” (Pyett, 2003). Critical reference groups are used in participatory action research and is a generic term for those the research and evaluation is intended primarily to benefit (Wadsworth, 1998). Subjecting the interview questions to evaluation by this kind of group may help to construct understandable questions that make better sense of the studied phenomenon by asking the “right questions in the right way.”

From the viewpoint of credibility, self-awareness of the researcher is essential (Koch, 1994). Pre-interviews may help to determine whether the interview questions are suitable for obtaining rich data that answer the proposed research questions. Interview tapes, videos, and transcribed text should be examined carefully to critically assess the researcher’s own actions. For instance, questions should be asked such as “Did I manipulate or lead the participant?” and “Did I ask too broad or structured questions?” Such evaluation should not only begin at the start of the study but also be supported by continuous reflection to ensure the trustworthiness of content analysis.

To manage the data, pre-testing of the analysis method is as important in qualitative as in quantitative research. When using a deductive content analysis approach, the categorization matrix also needs to be pretested in a pilot phase (Schreier, 2012). This is essential, especially when two or more researchers are involved in the coding. In trial coding, researchers independently try out the coding of the newly developed matrix (Schreier, 2012) and then discuss any apparent difficulties in using the matrix (Kyngäs et al., 2011) and the units of coding they have interpreted differently (Schreier, 2012). Based on their discussion, the categorization matrix is modified, if needed.

Sampling Strategy

From the viewpoint of sampling strategy, it is essential to ask questions such as the following: What is the best sampling method for my study? Who are the best informants for my study and what criteria to use for selecting the participants? Is my sample appropriate? Are my data well saturated? Thoroughness as a criterion of validity refers to the adequacy of the data and also depends on sound sampling and saturation (Whittemore, Chase, & Mandle, 2001). It is important to consider the sampling method used in qualitative studies (Creswell, 2013). Based on our research, the sampling method is rarely mentioned in qualitative content analysis

studies (Kyngäs et al., 2011). In qualitative research, the sampling strategy is usually chosen based on the methodology and topic, and not by the need for generalizability of the findings (Higginbottom, 2004). Types of qualitative sampling include convenience, purposive, theoretical, selective, within-case and snowball sampling (Creswell, 2013; Higginbottom, 2004; Polit & Beck, 2012). However, the sample must be appropriate and comprise participants who best represent or have knowledge of the research topic.

The most commonly used method in content analysis studies is purposive sampling (Kyngäs, Elo, Pölkki, Kääriäinen, & Kanste, 2011): purposive sampling is suitable for qualitative studies where the researcher is interested in informants who have the best knowledge concerning the research topic. When using purposeful sampling, decisions need to be made about who or what is sampled, what form the sampling should take, and how many people or sites need to be sampled (Creswell, 2013). However, a disadvantage of purposive sampling is that it can be difficult for the reader to judge the trustworthiness of sampling if full details are not provided. The researcher needs to determine which type of purposeful sampling would be best to use (Creswell, 2013), and a brief description of the sampling method should be provided.

Dependability refers to the stability of data over time and under different conditions. Therefore, it is important to state the principles and criteria used to select participants and detail the participants’ main characteristics so that the transferability of the results to other contexts can be assessed (e.g., see Moretti et al., 2011). The main question is then, “Would the findings of an inquiry be repeated if it were replicated with the same or similar participants in the same context (Lincoln & Guba, 1985; Polit & Beck, 2012)?” According to Lincoln and Guba’s (1985) criteria for establishing credibility, researchers must ensure that those participating in research are identified and described accurately. To gather credible data, different sampling methods may be required in different studies.

Selection of the most appropriate sample size is important for ensuring the credibility of content analysis study (Graneheim & Lundman, 2004). Information on the sample size is essential when evaluating whether the sample is appropriate. There is no commonly accepted sample size for qualitative studies because the optimal sample depends on the purpose of the study, research questions, and richness of the data. In qualitative content analysis, the homogeneity of the study participants or differences expected between groups are evaluated (Burmeister, 2012; Sandelowski, 1995a). For example, a study on the well-being and the supportive physical environment characteristics of home-dwelling elderly is likely to generate fairly heterogeneous data and may need more participants than if restrictions are applied, for example, studying only elderly aged above 85 years or those living in rural areas.

It has been suggested that saturation of data may indicate the optimal sample size (Guthrie et al., 2004; Sandelowski,

1995a). By definition, saturated data ensure replication in categories, which in turn verifies and ensures comprehension and completeness (Morse, Barrett, Mayan, Olson, & Spiers, 2002). If the saturation of data is incomplete, it may cause problems in data analysis and prevent items being linked together (Cavanagh, 1997). Well-saturated data facilitates its categorization and abstraction. It is easier to recognize when saturation is achieved if data are at least preliminarily collected and analyzed at the same time (Guthrie et al., 2004; Sandelowski, 1995a, 2001). It is common that all data are first collected and then analyzed later. We recommend that preliminary analysis should start, for example, after a few interviews. When saturation is not achieved, it is often difficult to group the data and create concepts (Elo & Kyngäs, 2008; Guthrie et al., 2004; Harwood & Garry, 2003), preventing a complete analysis and generating simplified results (Harwood & Garry, 2003; Weber, 1990).

Selection of a Suitable Unit of Analysis

The success of data collection should be assessed in relation to the specific research questions and study aim. The preparation phase also involves the selection of a suitable unit of analysis, which is also important for ensuring the credibility of content analysis. The meaning unit can, for example, be a letter, word, sentence portion of pages, or words (Robson, 1993). Too broad a unit of analysis will be difficult to manage and may have various meanings. Too narrow a meaning unit may result in fragmentation. The most suitable unit of analysis will be sufficiently large to be considered as a whole but small enough to be a relevant meaning unit during the analysis process. It is important to fully describe the meaning unit when reporting the analysis process so that readers can evaluate the trustworthiness of the analysis (Graneheim & Lundman, 2004). However, in previous scientific articles, the unit of analysis has often been inadequately described, making it difficult to evaluate how successful was the meaning unit used (Kyngäs et al., 2011).

Trustworthiness of Organization Phase in Content Analysis Study

According to Moretti et al. (2011), the advantage of qualitative research is the richness of the collected data and such data need to be interpreted and coded in a valid and reliable way. In the following sections, we discuss trustworthiness issues associated with the organization phase. In this phase, it is essential to consider whether the categories are well created, what the level of interpretation is, and how to check the trustworthiness of the analysis.

As part of the organization phase, an explanation of how the concepts or categories are created should be provided to indicate the trustworthiness of study. Describing the concepts and how they have been created can often be challenging, which may hinder a complete analysis, particularly if the

researcher has not abstracted the data, or too many different types of items have been grouped together (Dey, 1993; Hickey & Kipping, 1996). In addition, a large number of concepts usually indicates that the researcher has been unable to group the data, that is, the abstraction process is incomplete, and categories may also overlap (Kyngäs et al., 2011). In this case, the researcher must continue the grouping to identify any similarities within and differences between categories.

According to Graneheim and Lundman (2004), an essential consideration when discussing the trustworthiness of findings from a qualitative content analysis is that there is always some degree of interpretation when approaching a text. All researchers have to consider how to confirm the credibility and conformability of the organization phase. Conformability of findings means that the data accurately represent the information that the participants provided and the interpretations of those data are not invented by the inquirer (Polit & Beck, 2012). This is particularly important if the researcher decides to analyze the latent content (noticing silence, sighs, laughter, posture etc.) in addition to manifest content (Catanzaro, 1988; Robson, 1993) as it may result in over interpretation (Elo & Kyngäs, 2008). It is recommended that the analysis be performed by more than one person to increase the comprehensivity and provide sound interpretation of the data (Burla et al., 2008; Schreier, 2012). However, high intercoder reliability (ICR) is required when more than one coder is involved in deductive data analysis (Vaismoradi et al., 2013). Burla, Knierim, Barth, Duetz, and Abel (2008) have demonstrated how ICR assessment can be used to improve coding in qualitative content analysis. This is useful when using deductive content analysis, which is based on a categorization matrix or coding scheme.

However, there are no published recommendations on how the trustworthiness should be checked if the inductive content analysis is conducted by two or more researcher. Our suggestion is that one researcher is responsible for the analysis and others carefully follow-up on the whole analysis process and categorization. All the researchers should subsequently get together and discuss any divergent opinions concerning the categorization, like in the pilot phase mentioned earlier. For example, in one of our studies, two research team members checked the adequacy of the analysis and asked for possible complements (Kyngäs et al., 2011).

One study (Kyngäs et al., 2011) has suggested that data are most often analyzed by one researcher, especially when using inductive content analysis. In such a case, the credibility of the analysis can be confirmed by checking for the representativeness of the data as a whole (Thomas & Magilvy, 2011). According to Pyett (2003), a good qualitative researcher cannot avoid the time-consuming work of returning again and again to the data, to check whether the interpretation is true to the data and the features identified are corroborated by other interviews. Face validity has also been used to estimate the trustworthiness of studies (Cavanagh,

1997; Downe-Wamboldt, 1992; Hickey & Kipping, 1996). In this case, the results are presented to people familiar with the research topic, who then evaluate whether the results match reality. If the deductive approach is used, double-coding often helps to assess the quality of categorization matrix. According to Schreier (2012), if the code definitions are clear and subcategories do not overlap, then two rounds of independence coding should produce approximately the same results.

The value of dialogue among co-researchers has often been highlighted and it has been suggested that the participant's recognition of the findings can also be used to indicate the credibility or conformability (Graneheim & Lundman, 2004; Saldaña, 2011). However, it has been recommended that this be undertaken with caution (Ryan-Nicholls & Will, 2009). Some studies have used member checks, whereby participants check the research findings to make sure that they are true to their experiences (Holloway & Wheeler, 2010; Koch, 1994; Saldaña, 2011; Thomas & Magilvy, 2011). Although Lincoln and Guba (1985) have described member checks as a continuous process during data analysis (e.g., by asking participants about hypothetical situations), it has largely been interpreted and used by researchers for verification of the overall results with participants. Although it may seem attractive to return the results to the original participants for verification, it is not an established verification strategy. Several methodologists have warned against basing verification on whether readers, participants, or potential users of the research judge the analysis to be correct, stating that it is actually more often a threat to validity (Morse et al., 2002). Pyett (2003) has argued that the study participants do not always understand their own actions and motives, whereas researchers have more capacity and academic obligation to apply critical understanding to accounts.

Reporting Phase From the Viewpoint of Content Analysis Trustworthiness

Writing makes something disappear and then reappear in words. This is not always easy to achieve with rich data sets, as encountered in nursing science. The problem with writing is that phenomena that may escape all representation need to be accurately represented in words (van Manen, 2006). According to Holdford (2008), the analysis and reporting component of content analysis should aim to make sense of the findings for readers in a meaningful and useful way. However, little attention has been paid to the most important element of qualitative studies: the presentation of findings in the reports (Sandelowski & Leeman, 2011). In the next sections, we discuss trustworthiness issues associated with the reporting results, methods, and analysis process.

Reporting Results

Reporting results of content analysis is particularly linked to transferability, conformability, and credibility. Results

should be reported systematically and carefully, with particular attention paid to how connections between the data and results are reported. However, the reporting of results systematically can often be challenging (Kyngäs et al., 2011). Problems with reporting results can be a consequence of unsuccessful analysis (Dey, 1993; Elo & Kyngäs, 2008) or difficulties in describing the process of abstraction because it in part depends on the researcher's insight or intuitive action, which may be difficult to describe to others (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004).

The content and structure of concepts created by content analysis should be presented in a clear and understandable way. It is often useful to provide a figure to give an overview of the whole result. The aim of the study dictates what research phenomena are conceptualized through the analysis process. However, conception may have different objectives. For example, the aim of the study may be merely to identify concepts. In contrast, if the aim is to construct a model, the results should be presented as a model outlining the concepts, their hierarchy, and possible connections. Content analysis per se does not include a technique to connect concepts (Elo & Kyngäs, 2008; Harwood & Garry, 2003). The main consideration is to ensure that the structure of results is equivalent and answers the aim and research questions.

From the perspective of trustworthiness, the main question is, "How can the reader evaluate the transferability of the results?" Transferability refers to the extent to which the findings can be transferred to other settings or groups. (Koch, 1994; Polit & Beck, 2012). Authors may offer suggestions about transferability, but it is ultimately down to the reader's judgment as to whether or not the reported results are transferable to another context (Graneheim & Lundman, 2004). Again, this highlights the importance of ensuring high quality results and reporting of the analysis process. It is also valuable to give clear descriptions of the culture, context, selection, and characteristics of participants. Trustworthiness is increased if the results are presented in a way that allows the reader to look for alternative interpretations (Graneheim & Lundman, 2004). We fully agree with van Manen (2006) that qualitative methods require sensitive interpretive skills and creative talents from the researcher. Thus, scientific writing is a skill that needs to be enhanced by writing and comparing others' analysis results.

It has been argued that the use of quotations is necessary to indicate the trustworthiness of results (Polit & Beck, 2012; Sandelowski, 1995a). Conformability refers to objectivity and implies that the data accurately represent the information that the participants provided and interpretations of those data are not invented by the inquirer. The findings must reflect the participants' voice and conditions of the inquiry, and not the researcher's biases, motivations, or perspectives (Lincoln & Guba, 1985; Polit & Beck, 2012). This is one reason why authors often present representative quotations from transcribed text (Graneheim & Lundman, 2004), particularly to show a connection between the data and results. For example, each main concept should be linked to the data by a

quotation. Examples of quotations from as many participants as possible help confirm the connection between the results and data as well as the richness of data. However, the systematic use of quotations needs careful attention. Ideally, quotations should be selected that are at least connected to all main concepts and widely representative of the sample. However, there is a risk that quotations may be overused, thus weakening the analysis (Downe-Wamboldt, 1992; Graneheim & Lundman, 2004, Kyngäs et al., 2011). For example, if quotations are overused in the Results section, the results of the analysis may be unclear.

According to Hsieh and Shannon (2005), an important problem is failure to develop a complete understanding of the context, resulting in failure to identify the key categories. In such a case, findings do not accurately represent the data. To ensure the trustworthiness and especially credibility of the results, it is important to evaluate how well categories cover the data and identify whether there are similarities within and differences between categories. In addition, failure to complete the analysis abstraction process may mean that concepts are presented as results that are not mutually exclusive, leading to oversimplistic conclusions (Harwood & Garry, 2003; Weber, 1990). An incomplete analysis may involve the use of everyday expressions or repetition of respondents' statements and/or their opinions rather than reporting the results of the analysis (Kyngäs et al., 2011).

Reporting the Analysis Process

Without a full description of the analysis and logical use of concepts, it is impossible to evaluate how the results have been created and their trustworthiness (Guthrie et al., 2004). An accurate description of the analysis and the relationship between the results and original data allow readers to draw their own conclusions regarding the trustworthiness of the results. In nursing science, the number of methods concerning content analysis published in books and scientific articles has increased considerably over the last decade (Elo & Kyngäs, 2008; Harwood & Garry, 2003; Hsieh & Shannon, 2005; Neuendorf, 2002; Schreier, 2012). This may have led to improvements in the quality of reports on the process of content analysis. More attention is now paid to descriptions of the analysis, results, and how to evaluate the trustworthiness of studies. Consequently, this makes it easier for readers to evaluate the trustworthiness of studies.

The dependability of a study is high if another researcher can readily follow the decision trail used by the initial researcher (Thomas & Magilvy, 2011). Whitemore et al. (2001) have argued that vividness involves the presentation of rich, vivid, faithful, and artful descriptions that highlight the salient themes in the data. The analysis process should be reported in an adequate manner regardless of the methods used to present the findings (see Moretti et al., 2011). Steps should be taken to demonstrate credibility in research reports to ensure the trustworthiness of the content analysis. Monograph research reports facilitate detailed descriptions

of the analysis process and the use of figures, tables, and attachments to explain the categorization process. Based on our experiences, evaluation of the trustworthiness of results as a reader can often be difficult because of insufficient description of the analysis process (Kyngäs et al., 2011). Journal articles generally focus on the results rather than describing the content analysis process. All too often, the use of qualitative content analysis is only briefly mentioned in the methodology section, making it hard for readers to evaluate the process. A key question is, "In what detail should trustworthiness be presented in scientific articles?"—particularly as word limits often apply.

The fact that pictures may convey results more clearly than words should be borne in mind when reporting content analysis findings. The use of figures can be highly effective when reporting content analysis findings, especially when explaining the purpose and process of the analysis and structure of concepts. Very often, these aspects can be shown in the same figure, for example, a diagram that illustrates the hierarchy of concepts or categories may also give an insight into the analysis process (see, for example, Timlin, Riala, & Kyngäs, 2013). After reporting the results, a discussion of the trustworthiness of the analysis should be provided. It should be based on a defined set of criteria that are followed logically for each qualitative content analysis phase.

Discussion

The main purpose of this article was to discuss and highlight factors affecting trustworthiness of qualitative content analysis studies. The literature review used here was not a systematic review, so there are some limitations. First, we recognize that this is not a full description of trustworthiness and some points may be missing. For example, the language restrictions may have influenced the findings; research studies in other languages might have added new information to our description. Further studies are needed to systematically evaluate the reporting of content analysis in scientific journals, that is, to examine what researchers have emphasized when reporting the trustworthiness of their qualitative content analysis study, and how criteria of trustworthiness have been interpreted by those studies. This may help to develop a more complete description of trustworthiness in qualitative content analysis. However, the present methodological article was written by several authors who have extensive experience in using the content analysis method. In addition, the authors' experience as researchers, teachers, and supervisors of master's and doctoral students lends weight to our discussion.

Holloway and Wheeler (2010) have stated that researchers often have difficulty in agreeing on how to judge the trustworthiness of their qualitative study. The aim of this article was to identify factors affecting qualitative content analysis trustworthiness from the viewpoint of data collection and reporting of results. Qualitative researchers are advised to be systematic and well organized to enhance the

trustworthiness of their study (Saldaña, 2011). According to Schreier (2012), content analysis is systematic because all relevant material is taken into account, a sequence of steps is followed during the analysis, and the researcher has to check the coding for consistency. The information presented here raises important issues about the use and development of content analysis. If the method is thoroughly documented for all phases of the process (preparation, organization, and reporting), all aspects of the trustworthiness criteria are increased.

Before choosing an analysis method, the researcher should select the most suitable method for answering the target research question and consider whether the data richness is sufficient for using content analysis. Prior to using the method, the researcher should ask the question, "Is this method the best available to answer the target research question?" No analysis method is without drawbacks, but each may be good for a certain purpose. It is essential for researchers to delineate the approach they are going to use to perform content analysis before beginning the data analysis because the use of a robust analytic procedure will increase the trustworthiness of the study (Hsieh & Shannon, 2005).

Qualitative content analysis is a popular method for analyzing written material. This means that results spanning a wide range of qualities have been obtained using the method. Content analysis is a methodology that requires researchers who use it to make a strong case for the trustworthiness of their data (Potter & Levine-Donnerstein, 1999; Sandelowski, 1995a). Every finding should be as trustworthy as possible, and the study must be evaluated in relation to the procedures used to generate the findings (Graneheim & Lundman, 2004). In many studies, content analysis has been used to analyze answers to open-ended questions in questionnaires (Kygäs et al., 2011). However, such answers are often so brief that it is difficult to use content analysis effectively; reduction, grouping, and abstraction require rich data. In addition, trustworthiness has often been difficult to evaluate because articles have mainly focused on reporting the analysis of quantitative rather than qualitative data obtained in the study. Whether this affects the trustworthiness of the results can only be speculated upon. However, if researchers use content analysis to analyze answers to open-ended questions, they should provide an adequate description so that readers are able to readily evaluate its trustworthiness. Content analysis has also been commonly used in quantitative studies to analyze answers to open-ended questions.

There is a need for a self-criticism and good analysis skills when conducting qualitative content analysis. Any qualitative analysis should include continuous reflection and self-criticism by the researcher (Pyett, 2003; Thomas & Magilvy, 2011) from the beginning of the study. The researcher's individual attributes and perspectives can have an important influence on the analysis process (Whittemore et al., 2001). It is possible to obtain simplistic results using any method even when analysis skills are

lacking (Weber, 1990). According to Neuendorf (2002), the content analysis method can be as easy or as difficult as the researcher allows. Many researchers still perceive it as a simple method, and hence, it is widely used. However, inexperienced researchers may be unable to perform an accurate analysis because they do not have the knowledge and skills required. This can affect the authenticity (Lincoln & Guba, 1985; Whittemore et al., 2001) of the study, which refers to the extent to which researchers fairly and faithfully show a range of realities. A simplified result may be obtained if the researcher is unable to use and report the results correctly.

Furthermore, the reporting of the content analysis process should be based on self-critical thinking at each phase of the analysis. Whittemore et al. (2001) have argued that integrity is demonstrated by ongoing self-reflection and self-scrutiny to ensure that interpretations are valid and grounded in the data. Not only should a sufficient description of the analysis be provided to help validate the data, but the researcher should also openly discuss the limitations of the study. We agree with Creswell's (2013) comment that validation in a qualitative study is an attempt to assess the accuracy of the findings, as best described by the researcher and the participants. This means that any report of research is a representation by the author. Discussion of the trustworthiness of a study should be based on a defined set of criteria that are followed logically. Although many criteria have been proposed to evaluate the trustworthiness of qualitative studies, they have rarely been followed. It is recommended that authors clearly define their validation terms (see example from Tucker, van Zandvoort, Burke, & Irwin, 2011) because there are many types of qualitative validation terms in use, for example, trustworthiness, verification, and authenticity (Creswell, 2013).

Conclusion

The trustworthiness of content analysis results depends on the availability of rich, appropriate, and well-saturated data. Therefore, data collection, analysis, and result reporting go hand in hand. Improving the trustworthiness of content analysis begins with thorough preparation prior to the study and requires advanced skills in data gathering, content analysis, trustworthiness discussion, and result reporting. The trustworthiness of data collection can be verified by providing precise details of the sampling method and participants' descriptions. Here, we showed how content analysis can be reported in a valid and understandable manner, which we anticipate will be of benefit to both writers and reviewers of scientific articles. As important qualitative research results are often reported as monograph reports, there is a need for further study to analyze published articles where content analysis is used. This may produce further information that helps content analysis writers present their studies in a more effective way.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

- Baxter, J. (2009). Content analysis. In R. Kitchin & N. Thrift (Eds.), *International encyclopedia of human geography* (Vol. 1, pp. 275-280). Oxford, UK: Elsevier.
- Burla, L., Knierim, B., Barth, K. L., Duetz, M., & Abel, T. (2008). From the text to coding: Intercoder reliability assessment in qualitative content analysis. *Nursing Research, 57*, 113-117.
- Burmeister, E. (2012). Sample size: How many is enough? *Australian Critical Care, 25*, 271-274.
- Catanzaro, M. (1988). Using qualitative analytical techniques. In P. Woods & M. Catanzaro (Eds.), *Nursing research: Theory and practice* (pp. 437-456). New York, NY: Mosby.
- Cavanagh, S. (1997). Content analysis: Concepts, methods and applications. *Nurse Researcher, 4*, 5-16.
- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks, CA: Sage.
- de Casterlé, B. D., Gastmans, C., Bryon, E., & Denier, Y. (2012). QUAGOL: A guide for qualitative data analysis. *International Journal of Nursing Studies, 49*, 360-371.
- Dey, I. (1993). *Qualitative data analysis: A user-friendly guide for social scientists*. London, England: Routledge.
- Downe-Wamboldt, B. (1992). Content analysis: Method, applications and issues. *Health Care for Women International, 13*, 313-321.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*, 107-115.
- Emden, C., Hancock, H., Schubert, S., & Darbyshire, P. (2001). A web of intrigue: The search for quality in qualitative research. *Nurse Education in Practice, 1*, 204-211.
- Emden, C., & Sandelowski, M. (1999). The good, the bad and the relative, part two: Goodness and the criterion problem in qualitative research. *International Journal of Nursing Practice, 5*, 2-7.
- Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today, 24*, 105-112.
- Guthrie, J., Yongvanich, K., & Ricceri, F. (2004). Using content analysis as a research method to inquire into intellectual capital reporting. *Journal of Intellectual Capital, 5*, 282-293.
- Harwood, T. G., & Garry, T. (2003). An overview of content analysis. *The Marketing Review, 3*, 479-498.
- Hickey, G., & Kipping, E. (1996). A multi-stage approach to the coding of data from open-ended questions. *Nurse Researcher, 4*, 81-91.
- Higginbottom, G. M. (2004). Sampling issues in qualitative research. *Nurse Researcher, 12*, 7-19.
- Holdford, D. (2008). Content analysis methods for conducting research in social and administrative pharmacy. *Research in Social & Administrative Pharmacy, 4*, 173-181.
- Holloway, I., & Wheeler, S. (2010). *Qualitative research in nursing and healthcare*. Oxford, UK: Blackwell.
- Hsieh, H.-F., & Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*, 1277-1288.
- Koch, T. (1994). Establishing rigour in qualitative research: The decision trail. *Journal of Advanced Nursing, 19*, 976-986.
- Koch, T., & Harrington, A. (1998). Reconceptualizing rigour: The case for reflexivity. *Journal of Advanced Nursing, 28*, 882-890.
- Kyngäs, H., Elo, S., Pölkki, T., Kääriäinen, M., & Kanste, O. (2011). Sisällönanalyysi suomalaisessa hoitotieteellisessä tutkimuksessa [The use of content analysis in Finnish nursing science research]. *Hoitotiede, 23*(2), 138-148.
- Lincoln, S. Y., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.
- Moretti, F., van Vliet, L., Bensing, J., Deledda, G., Mazzi, M., Rimondini, M., . . . Fletcher, I. (2011). A standardized approach to qualitative content analysis of focus group discussions from different countries. *Patient Education & Counseling, 82*, 420-428.
- Morgan, D. L. (1993). Qualitative content analysis: A guide to paths not taken. *Qualitative Health Research, 1*, 112-121.
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods, 1*(2), 1-19.
- Neuendorf, K. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage.
- Neuendorf, K. (2011). Content analysis—A methodological primer for gender research. *Sex Roles, 64*, 276-289.
- Polit, D. F., & Beck, C. T. (2012). *Nursing research: Principles and methods*. Philadelphia, PA: Lippincott Williams & Wilkins.
- Potter, J. W., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research, 27*, 258-284.
- Pyett, P. M. (2003). Validation of qualitative research in the “real world.” *Qualitative Health Research, 13*, 1170-1179.
- Robson, C. (1993). *Real world research: A resource for social scientists and practitioner-researchers*. Oxford, UK: Blackwell.
- Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research & Development, 52*, 5-18.
- Ryan-Nicholls, K., & Will, C. (2009). Rigour in qualitative research: Mechanisms for control. *Nurse Researcher, 16*, 70-82.
- Saldaña, J. (2011). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage.
- Sandelowski, M. (1995a). Qualitative analysis: What it is and how to begin? *Research in Nursing & Health, 18*, 371-375.
- Sandelowski, M. (1995b). Sample size in qualitative research. *Research in Nursing & Health, 18*, 179-183.
- Sandelowski, M. (2001). Real qualitative researchers do not count: The use of numbers in qualitative research. *Research in Nursing & Health, 24*, 230-240.
- Sandelowski, M., & Leeman, J. (2011). Writing usable qualitative health research findings. *Qualitative Health Research, 22*, 1404-1413.
- Schreier, M. (2012). *Qualitative content analysis in practice*. Thousand Oaks, CA: Sage.
- Thomas, E., & Magilvy, J. K. (2011). Qualitative rigour or research validity in qualitative research. *Journal for Specialists in Pediatric Nursing, 16*, 151-155.
- Timlin, U., Riala, K., & Kyngäs, H. (2013). Adherence to treatment among adolescents in a psychiatric ward. *Journal of Clinical Nursing, 22*, 1332-1342.

- Tucker, P., van Zandvoort, M. M., Burke, S. M., & Irwin, J. D. (2011). The influence of parents and the home environment on preschoolers' physical activity behaviours: A qualitative investigation of childcare providers' perspectives. *BMC Public Health, 11*, Article 168.
- Vaismoradi, M., Bondas, T., & Turunen, H. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Journal of Nursing & Health Sciences, 15*, 398-405.
- van Manen, M. (2006). Writing qualitatively, or the demands of writing. *Qualitative Health Research, 16*, 713-722.
- Wadsworth, Y. (1998). *What is participatory action research? Action research international* (Paper 2). Retrieved from <http://www.aral.com.au/ari/p-ywadsworth98.html>
- Warr, D., & Pyett, P. (1999). Difficult relations: Sex work, love and intimacy. *Sociology of Health & Illness, 21*, 290-309.
- Weber, R. P. (1990). *Basic content analysis*. Newbury Park, CA: Sage.
- Whittemore, R., Chase, S. K., & Mandle, C. L. (2001). Validity in qualitative research. *Qualitative Health Research, 11*, 522-537.

Author Biographies

Satu Elo, PhD, is a Senior university lecturer in University of Oulu, Institute of Health Sciences. She is the second chairman of Finnish Research Society of Nursing Science. Her research and teaching area are both focusing on elderly care environments, and research methods especially from the viewpoint of theory development.

Maria Kääriäinen is Professor in University of Oulu, Institute of Health Sciences. She is in charge of the Teacher Education Program

in Health Sciences. Her research work has focused on two fields: 1) Health promotive counselling of chronically ill patients, and people with overweight, and 2) Effectiveness of education on the competence of nursing staff, students and teachers.

Outi Kanste (PhD) is a senior researcher at the National Institute for Health and Welfare in Finland. She has also worked at the University of Oulu and development projects of social and health services in municipalities. Her research interests are nursing leadership and management as well as service system and integration particularly in services for children, youth and families.

Tarja Pölkki, PhD, is an Adjunct Professor, senior researcher and lecturer in the Institute of Health Sciences, University of Oulu. Her research interests concern the methodological issues in nursing science, and the well-being of children and their families focusing on the aspects of pain assessment and non-pharmacological interventions, and promoting of child-and family-centeredness in nursing.

Kati Utriainen, PhD, is a Coordinator in University of Oulu, Institute of Health Sciences, Finland. She is active in conducting and developing web-based learning of physicians specializing in occupational health care and education of their trainer doctors. She also works as an occupational health nurse in the occupational health centre of the City of Oulu.

Helvi Kyngäs is Professor in University of Oulu, Institute of Health Sciences. She is also head on Nursing Science studies and head of PhD studies in Health Sciences. She is also working a Part-time Chief Nursing Officer at Northern Ostrobothnia Hospital.