THEMATIC ANALYSIS: A CRITICAL REVIEW OF ITS PROCESS AND EVALUATION

Alhojailan, Mohammed Ibrahim

PhD at UK/ DMU/ School of Technology/CCSR / Lecturer at King Saud University/ Educational Technology
Malhojailan@ksu.edu.sa

ABSTRACT

This paper critically reviews of the use of thematic analysis (TA) in qualitative research by describing its procedures and processes and by comparing grounded theory (GTA) with hermeneutic analysis. The literature that relates to thematic analysis (TA) shows that there is a lack of descriptions issues exist due respect to the concepts, process, validations and clarifications that been used by researchers.

This paper finds that thematic analysis is a comprehensive process where researchers are able to identify numerous cross-references between the data the research’s evolving themes (Hayes 1997). It provides flexibility for approaching research patterns in two ways, i.e. inductive and deductive (Frith and Gleeson 2004; Hayes 1997; Niece 2011; Halldorson 2009). This makes the process of thematic analysis more appropriate for analyzing the data when the research’s aim is to extract information to determine the relationship between variables and to compare different sets of evidence that pertain to different situations in same study.

Key words

Qualitative research, thematic analysis, validation, reliability, inductive and deductive research

1. Introduction

Qualitative data collection is usually dependent on interpretation. This means that the data requires several explanations. This is because huge amounts of qualitative evidence are often collected. Additionally, there is no distinction between data collection and its analysis (Cassell and Symon 1994), as supported by Cohen et al. (2011), who said that data analysis in qualitative research is distinguished by,

“Merging of analysis and interpretation and often by the merging of data collection with data analysis.” (p. 537)

This means that there is an overlap of analysis and interpretation to reach a conclusion.
In terms of qualitative data analysis, Denscombe (2010) mentioned few principles by for qualitative data analysis, he claims that by following them will probably result in more efficient outcomes. The first principle is to compact extensive and diverse raw data into a succinct structure. It could achieve by organising oral and write the data into charts and tables. This provided the researcher the opportunity to identify, compare and determine the data upon which to focus. The second principle is to make the relationship between the research objectives and the summary clear. That mostly fit when the objectives of any qualitative study considered the clear drivers responsible for its research and analytical methodologies. The third principle suggests that one should conclude by developing a model and/or improving the conceptual basis of the research.

Some researchers utilize programming for preparing and instructing the data, while others prefer to use traditional manual methods. In some instances, it may be better but not always to use manual analysis rather than computer-based methods, e.g. NVivo. Software is usefully able to analyse qualitative data in terms of gathering all the evidence and subsequently organising and grouping it into similar themes or ideas. In this regard, using software for analysing qualitative data is valuable in terms of improving the rigours of the analytical steps for validating that which does not reflect the researcher’s impressions of the data. Furthermore, software allows the researcher to analyse the data at a more specific level. Sometimes, however, software is less helpful. Welsh (2002) argued that software might not prove as helpful as one may expect. He said,

“In term of addressing issues of validity and reliability in thematic ideas that emerge during the data analysis process and this is due to the fluid and creative way in which these themes emerge.” (p. unknown)

Software requires not only careful labelling but also an organised system of filing the data in order to ensure that it is removed from files that are easily set in a new text, i.e. ‘recontextualised’.

2. Grounded Theory, Hermeneutic and Thematic Analysis, Concepts and Suitability

Braun and Clarke (2006) argue that Grounded Theory is very similar to Thematic Analysis in terms their procedures for coding ‘themes’ or coding from data (pp. 8-10). Despite these similarities, differences exist. These differences derive from a major feature of Grounded Theory. It has been suggested that their data collection and analysis processes run parallel. In other words, the data analysis process starts at the same time as the data collection process, which means that further data collection should be grounded on what has been previously analysed (Strauss and Corbin 1990). This approach is unsuitable for the researchers who seek to compare two separate sets of data that are gathered at different times. For example, if the aim of the researcher is to measure what differences occur among participants during the project but the data needs to be collected during two separate phases, (i.e. at the beginning and the end of the project.) then method of collecting pre- and post-data is not incompatible with thematic analysis (Miles and Huberman 1994).

Moreover, Grounded Theory analysis data that is undetermined before starting a study, e.g. if the sample used was determined and defined then Thematic and Hermeneutic Analysis more appropriate. Grounded Theory analysis relies on theoretical sampling, which is determined during data collection (Glaser and Strauss 1967).
Hermeneutic and Thematic Analysis are similar in that both focus on interpreting the data. Both are suitable to generate theory (Bryman 2008; Miles and Huberman 1994; Myers 1997; Hayes 2000). Hermeneutic Analysis, however, with respect to its analytical principle is partly referred to as,

“The dialectic between the understanding of the text as a whole and the interpretation of its parts, in which descriptions are guided by anticipated explanations.” (Myers 2004 p.107)

Boland (1985) argues that Hermeneutic Analysis can be best used to understand the organisation or the institution as a whole. This means that understanding a part (personal) will result in the understanding of the whole (organisation) and vice versa. In other words, Hermeneutic Analysis focuses on a wider research context and takes into account the entire background environment of the research during data collection (Myers 2004; Boland 1985). Hermeneutic Analysis is unlikely to be appropriate for analysing data that focuses only on participants’ visions due to issues.

3. THEMATIC ANALYSIS

Thematic Analysis is a type of qualitative analysis. It is used to analyse classifications and present themes (patterns) that relate to the data. It illustrates the data in great detail and deals with diverse subjects via interpretations (Boyatzis 1998).

Thematic Analysis is considered the most appropriate for any study that seeks to discover using interpretations. It provides a systematic element to data analysis. It allows the researcher to associate an analysis of the frequency of a theme with one of the whole content. This will confer accuracy and intricacy and enhance the research’s whole meaning. Qualitative research requires understanding and collecting diverse aspects and data. Thematic Analysis gives an opportunity to understand the potential of any issue more widely (Marks and Yardley 2004). Namey et al. (2008) said,

“Thematic Moves beyond counting explicit words or phrases and focuses on identifying and describing both implicit and explicit ideas. Codes developed for ideas or themes are then applied or linked to raw data as summary markers for later analysis, which may include comparing the relative frequencies of themes or topics within a data set, looking for code co-occurrence, or graphically displaying code relationships.” (p.138)

Thematic Analysis allows the researcher to determine precisely the relationships between concepts and compare them with the replicated data. By using, thematic analysis there is the possibility to link the various concepts and opinions of the learners and compare these with the data that has been gathered in different situations at different times during the project. All possibilities for interpretation are possible.
4. WHEN I CAN CHOOSE THEMATIC ANALYSIS

Thematic analysis is appropriate for the following situations

I. Data Interpretation

Firstly, good qualitative research needs to be able to draw interpretations and be consistent with the data that is collected. With this in mind, Thematic Analysis is capable to detect and identify, e.g. factors or variables that influence any issue generated by the participants. Therefore, the participants’ interpretations are significant in terms of giving the most appropriate explanations for their behaviours, actions and thoughts. This fits in well with the features that are involved in the process of Thematic Analysis (Hatch 2002; Creswell 2003).

II. Deductive and inductive approaches

Secondly, the flexibility of Thematic Analysis allows it to be used in both inductive and deductive methodologies (Frith and Gleeson 2004; Hayes 1997). For example, using an inductive approach the majority of the data that is collected will start with a precise content and then move to broader generalisations and finally to theories. This tends to ensure the themes are effectively linked to the data (Patton, 1990). This flexibility would enable any researcher to deal with the observational data that was collected throughout the study. In such a situation there needs to be a variety of options to frame the analysis around precise words, i.e. ‘explanations’, since collecting different data, e.g. observational data, attempts to give explanations for the study’s entire proceedings. Furthermore, for comparing the data collected with the perceptions of the participants other comparative methodologies have to be employed, e.g. the questionnaire as a deductive approach.

In addition, to investigate the adapted observational data in terms of the constructed meanings that emanated from participants’ opinions or feedback e.g. during the study, it could be achieved by an in-depth analysis with the main focus either on the perspectives of separate or groups of individuals (Hatch 2002; Creswell 2003).

III. Analysis two different phased of data

Thirdly, Thematic Analysis could be appropriate when the study aims to understand the current practices of any individual. In particular the influence of any variable, which is utilised by participants in a practical way in order to investigate and identify how current situations are influenced by their points of view. This approach fits in with analysing the different phases of data collection, e.g. pre-/post-data. Further, it works when the research seeks to examine the impact of the use of blogs on learners’ attitudes. For example, Alhojailan (2012), adapted a questionnaire to collect data for his study during two different phases, i.e. before and after the implementation of web 2.0 tools with learners. This approach made Thematic Analysis fit in this type of data collection. Thematic analysis is association with the data that was gathered in this research. It, therefore, highlights the flexibility of this approach, which enables it to cope when data has to be collected at different times – separately (Miles and Huberman 1994).

In addition, these two phases will allow the researcher to observe the differences and similarities that take place before and after data adaptation. With this in mind, Thematic Analysis is considered suitable to deal with this type of
data, through which the researcher can highlight the differences and similarities apparent within the data set (Creswell 2009; Creswell 2009; Boyatzis 1998).

IV. Coding and categorising

Finally, Thematic Analysis provides the opportunity to code and categorise data into themes. For example, how issues influence the perceptions of participants. In the case of Thematic Analysis, processed data can be displayed and classified according to its similarities and differences (Miles and Huberman 1994).

In order to achieve the above, the process should include coding, categorisation and noting patterns, i.e. different level of themes could be provide (Braun and Clarke 2006), also to provide a relationship between the variables and factors in order to create a reasonable and logical chain of evidence (Creswell 2009; Braun and Clarke 2006; Miles and Huberman 1994). By gathering data using different instruments, (e.g. observation, questionnaires with interviews on one study) with participants in different environments, Thematic Analysis will produce and present the data more effectively and reflect the reality of the data collection (Miles and Huberman 1994; Creswell 2009; Hayes 1997).

5. THEMATIC ANALYSIS MODEL

This section describes the Miles & Huberman (1994) model for the thematic analysis process. It consists of three link stages or ‘streams’, i.e. data reduction, data display and data conclusion-drawing/verifying as illustrated by the following figure.

Data reduction refers to the process of choosing, focusing, simplifying, building and transforming data (Miles & Huberman, 1994). During this stage, new thoughts and ideas are developed in terms of what should be included in the data display. Data display is described by Miles & Huberman (1994) as,
“An organised, compressed, assembly of information that permits conclusion drawing and action” (p. 11)

Importantly these stages focus on visualising the data by using a number of different display techniques, such as, quotations, narrative text, figures, tabulating differences and similarities and clarifying the relationship including its associated complexity of data (Miles and Huberman 1994; Gibbs 2002; Yin 2010).

The advantage of utilising different data display techniques makes the description of the comparison and similarities clearer, e.g. tabulating. In addition, it also increases the overall reliability of the research to make it valid for other researchers. Presenting different quotations is aimed to provide evidence, support and validate interpretations (Miles and Huberman 1994; Gibbs 2002; Patton, 1990).

The final stages of the data analysis process are linked by arranging and organising the research’s concepts and thoughts. This is achieved by building coherent findings and drawing structures of the results from the data that is displayed. During this stage, the meaning of contradictory and identical data needs clarification (Creswell, 2007; Miles & Huberman, 1994).

6. THEMATIC PROCESS

Before continuing to implement the data using Thematic Analysis, a few points that are worthwhile mentioning clearly show the basis for adapting thematic analysis to a specific piece of research. It highlights the aim of validation to establish a fundamental view of the analysis itself. This will assist the researchers with the details and sequence of how the raw data could be utilised.

If the data collection gathered based on e.g. observation that rely on participants’ visions - feeling or attitude - In Thematic Analysis data investigation and generating-theory are combined with its analytical element (Braun and Clarke 2006; Crawford et al. 2008). This is particularly appropriate when the researcher aims to examine the data in order to discover common themes and thoughts from more than one participant. It is beneficial to allocate a narrative to the diverse data to gain a clear logical understanding of the participant’s thoughts and to convey their experience (Crawford et al. 2008).

With same example, a better understanding of participants’ attitudes and reflections on issues which could be - mostly- measure is best gained through their diverse statements. Thematic Analysis provides the opportunity for researchers to move beyond calculating unambiguous words or statements or expressing the ideas. The themes develop the clues and then adapt or connect them to the raw data as summary indicators for deferred analysis. Namey & Namey et al. (2008) endorse this view and feel that by following these concepts thematic analysis,
“May include comparing the relative frequencies of themes or topics within a data set, looking for code co-occurrence or graphically displaying code relationships.” (p.138)

To present content when Thematic Analysis is used, the theme must “describe the bulk of the data” (Joffe and Yardley 2004 p.67). In other words, a large amount of content, i.e. data is required. This is because, while one single statement is significant, it does not necessarily reflect the full story. This is especially true when the research’s objectives aim to gain an insight and discover relationships between the diverse data that originated from the different groups of learners. Thus, the researcher needs to provide and describe a large amount of data (Ibid, pp.67–77).

Thematic Analysis provides rich, detailed and complex data, which is compatible with Braun & Clarke’s (2006) vision. Further, Blacker (2009) argues that a rich thematic description of the entire data would assist him and/or the readers to get a sense of “the predominant and important themes” (p.83) from the data.

A thematic analysis process analyses the data without engaging pre-existing themes, which means that it can be adapted to any research that relies -only on- upon participants’ clarifications. In other words, each statement or idea contributes towards understanding the issues, which leads to an appreciation of the whole picture. This is because every statement is valid in understanding a single concept or ones shared with other statements. Concepts are thus constructed to give a full picture of the learner’s views and actions. Furthermore, presenting similarities and differences between the participants’ perspectives will assist the readers to obtain a global view (Joffe & Yardley 2004; Blacker 2009).

These guiding principles should to be considered by any research that uses a thematic process for its data analysis. The following sections describe in detail the process of data analysis using the stages of the Miles & Huberman (1994) Model, i.e. data reduction, data display and drawing the data to give an overview for the validation of the themes that extracted.

6.1 DATA REDUCTION

Data reduction is the first stage in data analysis according the Miles & Huberman (1994) model. In addition, data reduction is,

“A form of analysis that sharpens, sorts, focuses, discards, and organizes data in such a way that “final” conclusion can be drawn and verified.” (Miles & Huberman 1994, p.11)

It includes the process of selecting, simplifying and transforming the data. Miles & Huberman (1994) argue that reducing and transforming data in qualitative research can be achieved in different ways. It could be,
“...through selection, through summary or paraphrase, through being subsumed in larger pattern” (p.11).

The procedure of data reduction is performed in such a way that conclusions are drawn and verifications are completed. Furthermore, coding is involved by assigning table units to the data that could be collected from the participants whether it was a single statement or a longer answer.

The main purpose of coding in Thematic Analysis is to make connections between different parts of the data. Coding is derived from the participants’ responses, e.g. statements and reports and it categorises information with the aim of framing it as theoretical perceptions(Coffey and Atkinson 1996). Coding will allow the researcher to review the whole of the data by identifying its most significant meaning or to put it simply what is the data trying to say or tell us (Miles and Huberman 1994; Halldorson 2009; Coffey and Atkinson 1996).

The following section will discuss and describe the different phases that could involved in data reduction.

6.1.1 PHASES IN DATA REDUCTION

The data reduction could be achieved through three main phases. Each phase, however, reduced the data in different ways as follows.

I. First phase for data reduction:

After collected the data, the researchers should tabulate it using Microsoft Word prior to preparing and organizing the content of data. This meant that the data are ready to be analysed word-by-word, using the tables to show any significant patterns or themes(Miles and Huberman 1994; Halldorson 2009). In addition, Bogdan and Biklen (2007)argue that in Thematic Analysis, data ‘must’ be read at least twice (p.165) so that the researcher should “get a feel for the text by handling your [the] data multiple times.” (Ryan and Bernard 2003 p.11) Regarding this issue, Bernard (2000) mentioned an ocular scan method, which he argued is one of the best ways for the researcher to ‘hunt’ for themes and patterns in qualitative data (see Attard and Coulson 2012 p.501; Kim 2008 p.12).

From the researcher’s experience of analysing the data from two of his papers, which are entitled “The effectiveness of Weblogs as e-learning tools in higher education”) and, “Identification of learners’ attitudes regarding the implementation of read/write web, blog tools: a case study in higher education”). Reading the data a few times before and after identifying the themes and codes proved beneficial for following reasons.

1. It allowed the researcher to appreciate the full picture and make connections between the participants’ thoughts, ideas and the data collected through observations.
2. Reading prior to starting analysis allowed the researcher to identify and have more time to evaluate the data so preventing precipitous conclusions.
II. The second phase of data reduction

This phase involved highlighting the sentences from each participant that could be used, for example, e.g.to answer the study’s questions by taking ‘excerpts from the participant’s full text’. It has been advised that researchers should at all times keep an eye on the study’s questions during data collection and analysis (Halldorson 2009), which will assist the researcher to identify accurately ‘excerpts’ that relate to the research’s objectives. In addition, Ryan and Bernard (2003) said,

“We highly recommend pawing through texts and marking them up with different coloured highlighter pens.” (p.11)

By that, therefore, all the data should be highlighted to prepare it for the next phase.

III. The Third phase of data reduction

This phase involved using the highlighted sentences and then breaking the data into smaller segments or themes. These segments or themes referred to the sentences of a paragraph. This established the first themes from the data. Subsequently, the researcher then should to read the full content again in order to compare, contrast and/or search for missing information that had not appeared in the first level of the themes (Ryan and Bernard 2003).

The data should be developed further under the first level of themes, which enabled a better tabulation of the themes texts and it should be saved in a new Word document. This procedure will make the themes clearer and more understandable in terms of the researcher’s focus. The data should then be prepared and readied for identifying and classifying the second level of themes.

Before proceeding to identify the second level of themes i.e. ‘codes’, it is worthwhile discussing the themes’ validity, which must be conducted during these phase. The aim is to make sure that the excerpts from the first themes represented the whole text. This is discussed in the next section.

6.1.2 Themes’ reliability and validity

One important step in Thematic Analysis is that the ‘themes’ need to be evaluated to ensure they represent the whole of the text. Miles & Huberman (1994) said that validating themes in the early and late stages of data analysis is essential. It has been suggested that the researcher should involve an outside reviewer during this early stage to evaluate and identified themes. In other words, to test if the themes the researcher identified are compatible with the whole of the text or not. The researcher should subsequently involve an independent reviewer for his/her feedback. This will enable the researcher to compare the two sets of feedback (Miles & Huberman 1994). The main purpose of this procedure is to “build reliability in themes analysis coding” (Hosmer 2008 p.52). The researcher is now better informed of any conflicting results (if there are any) with respect to any themes that were added or removed by the outside and independent reviewers (Miles & Huberman 1994; Hosmer 2008).
Subsequently, the researcher should refer to the list of themes agreed with the outside reviewer and identify those excerpts made by participants that support each theme. Late checking and verification involved the independent reviewer to evaluate the overall themes, demonstrate and confirm the details of textual excerpts “similar to validity in positivistic terms” (Hosmer 2008, p.52). Miles & Huberman (1994) say that by including two i.e. outside/independent reviewers at two separate phases would ‘probably’ build a strong process for analytical credibility “similar to reliability from a positivistic perspective” (Hosmer 2008, p.52).

Validation must be done in the early stages of any research. The first level of themes must be validated during the first and second level of themes during the second phase.

In addition, from the researcher’s experience, validation is beneficial for the data, especially at the first level of themes as it provides accurate and reliable ones for the second level. Furthermore, it makes the data at the second level of themes less prone to errors and mistakes. This could be notice at the second level of the themes’ ‘codes’. Here the agreed codes of the independent and outside reviewers were more consensual and that there were fewer differences between them.

6.1.3 Data display

The second main step of the Miles & Huberman Model (1994) is data display. This step involves retrieving data using data display (Coffey and Atkinson 1996). It cannot be separated from data reduction because it complements the former (Miles & Huberman 1994). Data display is “the organized, compressed assembly of information” (Ibid p.11). It aims to make sense of the data that is collected. Data display organises data, helps to arrange concepts and the thoughts (Miles & Huberman 1994).

Following theme reduction, the researcher should review the research questions to identify any information that relates to similar concepts.

In addition, displaying the data serves a number of purposes such as:

1. The ability to view and enhance the data more clearly for the research
2. To avoid data overload during the process of analysis
3. Making sense of the data that has been collected by displaying related concepts from different statements (Miles & Huberman 1994; Halldorson 2009)

All the data that related to each question should to organize and present in order. This allowed the researcher to explore any differences, similarities and interrelationships by entering the data into conceptual clusters for analysis. Data display was used descriptively to gain conceptual coherence by collating items that related to each research question (Miles & Huberman 1994). Data reduction will be described in detail (Ibid 1994; Braun & Clarke 2006). Analysis will go further at this point as it will include interpretations of the concepts involved in the research topic to support participants’ statements by supplying relevant evidence (Boyatzis 1998; Blacker 2009).
Data may be displayed using a variety of techniques in order to facilitate its analysis. It better to include figures, tables, graphs, charts, maps of categories, narrative text and quotations (Yin 2010). Miles & Huberman (19940 consider that all that kinds of data display aim to,

“Assemble organised information into immediately an accessible and compact form so that the analyst can see what is happening and either draw justified conclusions or move on to the next step of analysis the display suggests may be useful.” (p.11)

Displaying the data in a variety of ways e.g. tables, figures and theme maps provides opportunities to gain an extra in-depth understanding of the data it. Each piece of data offers the researcher the ability to better explanation the data, e.g. tabulated files provide an appropriate and convenient technique to assist and make comparisons through different themesmaps (Gibbs 2002). In addition,direct quotations provide supportive meaning to the data’s interpretation for some statements (Patton 2002).

Utilizing different data display techniques and gradually framing it, enables the researcher to focus and organise his/her thoughts by linking and comparing the information to reach conclusions (Miles & Huberman 1994; Gibbs 2002).

6.1.4. Data drawing and conclusions

The third step of the Miles &HubermanModel comprises data drawing and conclusions (1994). These workers have suggested the use of somepoints to assist researchers to draw conclusions having displayed data in a variety of ways. Some of their ideasto generate meaning from the data were adopted by this research. These included:

1- The notation of any patterns or themes and the relevance of any statement especially if similar or contrasting
2- Grouping or establishing categories of ‘information that can go together’
3- Identifying interrelations among factors and variables
4- Building conceptual coherence and consistency, which at the end it should use to explore the validity of the findings so that they fit the theoretical framework of the study.

The stages of data drawing and display must not become separated from data reduction, because they are complementary. In addition, the stages involve drawing data and verification (Miles & Huberman 1994).

7. Conclusion

This paper has discussed the differences between the three main concepts of data analysis, i.e. GT, HA and TA. Further, it described in detail the model of Miles & Huberman (1994). It showed that Thematic Analysis approaches are appropriate when samples are determined and defined before proceeding with the study. In contrast, analysis by Grounded Theory relies on theoretical sampling, which is determined during data collection (Glaser & Strauss
1967). In addition, Thematic Analysis offers the flexibility for starting data analysis at any time during the project, where there is no association between the data gathered and the result of the process itself. More importantly it provides the flexibility for approaching research patterns in two ways, i.e. inductive and deductive (Frith and Gleeson 2004; Hayes 2000; Halldorson 2009).

Finally, in basic terms, Thematic Analysis provides a comprehensive process for a researcher to identify numerous cross-references between the evolving themes and the entire data (Hayes 1997). By using Thematic Analysis, it is possible to link the various concepts and opinions of participants and compare them with the data that has been gathered in different situation at different times from other or the same participants during the project. In this case, the potential for interpretation becomes infinite.

8. Acknowledgements

I would like to thank Professor Bernd Carsten Stahl and Mr. Ralph Birkenhead for their unlimited support. I also wish to thank Dr. Jehad Alamri for his valuable feedback and comments.

9. References


