Building risk matrices from interview transcripts utilising HCA and IPA

Megan Field¹

¹University of Birmingham

ABSTRACT

This paper proposes a methodological guideline for analysing interview transcripts to aid in the construction of risk matrices. This is to allow for the tabulation of qualitative data in a suitable manner as to provide appropriate qualitatively informed recommendations. Using this methodology, a comprehensive and qualitatively supported table to register concerns, priority and/or urgency of themes is created that can address inter- and intra- actor factors in socio-technical systems. The analysis aims to communicate the in-depth, rich data of narrative inquiry in verbal protocols to more technical or quantitative domains.

KEYWORDS

Interviews, Qualitative Research, System Ergonomics, Hierarchical Content Analysis, Interpretative Phenomenological Analysis

Introduction

Qualitative research methods are increasingly used to explore user or operator observations in human factors and ergonomics research. These involve the use of unstructured, semi-structured and structured verbal protocol, such as interviews and focus groups (Brinkmann, 2014). However, guidelines to analyse these narrative units are limited in the psycho-philosophical pedagogy (Howard-Payne, 2016) and within ergonomics model frameworks are the primary method of communicating complex organisational relationships between actants within a system. These theoretical frameworks exist in parallel within both perspectives; for example, the macro-ergonomic Socio-Technical System Model (Rasmussen, 1997) and the constructivist Conditional Matrix (Strauss & Corbin, 1990) both seek to describe the shifting dynamics between humans, objects and human behaviour processes within a multi-levelled system. In both, it is difficult to know how to appropriately populate the frameworks with information gathered from large data sets, such as interview transcriptions. One method is content analysis, based on grounded theory (Glaser & Strauss, 1967) (Glaser & Strauss, 2017). However, this technique does not explore the rich descriptive narrative inquiry useful for identifying underlying expressions in dominant themes. The importance of these are to identify barriers, faults or incidents in complex systems which may not be identified in a sweeping quantitative method such as content analysis. Therefore, this approach seeks to systematically analyse qualitative data through Interpretive Phenomenological Analysis (IPA) and Hierarchical Content Analysis (HCA). This is to allow us to appropriately extract information from interview scripts to build risk matrices, registers of concerns and other human factors reporting techniques, thus communicating actant or process issues in macro- and microergonomics systems.

Aims and Assumptions of the Analysis

The approach described in this paper assumes that the qualitative data collected is semi-structured, open-ended and the audio recordings from the verbal protocol of these interviews have been transcribed in full. The examples used in this paper are from a study exploring dominant illness narratives and expressions of trust of disabled graduates towards post-university employment recruitment processes. The qualitative data analysis computer software package, NVIVO (Version 10) (QSR International Pty Ltd, 2012), was used in the examples provided to codify and organise references. Leximancer (Leximancer Pty Ltd.) would also be suitable dependant on the researcher's preference and expertise. Note, the example data in this paper utilised responsive interviewing methodology (Rubin & Rubin, 2011) whereby conversational partnership between researcher and participant is encouraged and can be used to inform the semi-structured interviewing protocol for future or further interviews. The method discussed could also be adapted to analyse data sets from structured interview techniques.

The framework adopts a phenomenological methodology to explore the narrative inquiry of the transcripts and borrows methods from constructivist grounded theory (Charmaz, 2006) in order to ground any social theory in experiences. Qualitative research must accept that researcher bias exists and can be accommodated through transparency and acknowledgement in the interpretation of findings and theoretical framework underpinnings. The method was developed as an exploration of more homogenous data sets commonly seen in IPA compared to other sources of analysis of qualitative data (Creswell, 1998) as well as providing rich, contextual complexity to the clarification of themes.

The transcripts utilise interpretive phenomenological analysis for initial identification, coding of raw themes and clusters of data units and hierarchical content analysis for classification of categories, central themes and extrapolation of narratives. The analysis method seeks to systematically review themes to inform risk matrices and concern registers at both the sematic and latent thematic level (Braun, Clarke, & Terry, 2014). The use of interpretative and descriptive approaches to classify the data provides a method of numerical communication of qualitative research to help convey recommendations in technical (and/or quantitative reliant) domains and to appropriately communicate the in-depth rich data of narrative inquiry.

Stages of the Analysis

Stage 1 – Notetaking and Preliminary Categorisation

The researcher should make notes during and post-interview regarding the areas and topics discussed throughout the interview. This is more streamlined for those who undertake semi-structured and structured interview protocol. These 'memos' (Pandit, 1996) will become initiators which will illicit how to categorise data, record ideas and aid in the development of themes in later stages.

Stage 2 – Immersion

At this stage, the researcher gathers in-depth familiarity of the data sets through immersion in the data. Through re-reading the transcripts, the researcher can identify major points of interest, general feel and context which will generate exploratory codes.

Stage 3 – Interpretive Phenomenological Analysis

The analysis process of Interpretative Phenomenological Analysis and Hierarchical Content Analysis are similar in the first three phases (see Table *1* and Table *2*). In this method, the similarity provides parallel analysis in the exploration of the transcripts. This will influence the raw themes generated for the hierarchical content analysis. IPA is used as it facilitates the immersion required to adequately analyse small or homogenous data sets such as those which use interview data (whether for narrative inquiry, user experience, focus groups and more). Furthermore, IPA seeks to describe and explore lived experiences, thus is appropriate for exploring narratives as it reports rich descriptions and characteristics identified in the data and framing the context.

The immersion in the data of this stage provides the context and knowledge of the data set which will shape the categorisation of the risk matrix consequences in stage seven.

Table 1 - Analysis Procedure for Interpretative Phenomenological Analysis (Adapted from Sparkes and Smith (2013))

Phase		Description of Process				
1.	Immersion/Searching for	In-depth familiarity of data through engagement with				
	themes in the first case	collected information (e.g. multiple re-reading of				
		data set transcriptions). Exploratory codes generated				
	from first case.					
2.	Identifying and Labelling	Data sets are coded to capture essential quality of				
	Themes	transcript.				
3.	Connecting Theme	Refocus analysis to identify themes from data.				
		Cluster and connect emerging concepts that share				
		similar meaning.				

Stage 4 – Hierarchical Content Analysis

Hierarchical content analysis provides secondary analysis and seeks to explore themes intertranscript through systematic cross-checking and confirmation. In this approach, HCA is used to outline the context in which the weighting of the themes (as described in Stage 5 and 6) is based on. The exploratory themes outlined in the previous stage are then clustered into related units (coded nodes) in a clear and concise tabulated form.

Table 2 - Analysis Procedure for Hierarchical Content Analysis (Adapted from Sparkes and Smith (2013)

Phase	Description of Process		
4. Cross-checking	Examination of identified themes and clustered groupings against original data set.		
5. Confirmation	Review of raw themes and clustered groupings.		
6. Produce a Table	Production of data within a table of hierarchical structure to display units generated by nature of themes.		



Due to the diversity and large size of transcription data sets, it can be difficult to objectively explore which themes are more prevalent than others, and whether they are relevant compared to their super-ordinate themes. Therefore, this approach seeks to succinctly divide salient themes into 'absolute' and 'relevant' themes. The term 'absolute' in this approach refers to the amount of words coded in references to the super- and sub- ordinate themes occurring in the total word count of the data set. For example (see Table 3), the number of references discussing all topics relating to culture equal 27685, which equates to 81.7% of the entire interview data set. However, to address the relevancy of these sub-themes within the category boundaries, the relative content weight was calculated. As references can be coded to multiple themes, the node summary report provided by NVIVO (QSR International Pty Ltd, 2012) provides number of words coded to reference nodes (total words referenced at all nodes is equal to 32.8%. Further breakdowns of salient themes are provided in Table 5.

Table 3 - Table Outlining the Calculation of Absolute and Relative Theme Weighting (Example)

	Document Word Count	33888	
	Coded Word Count	84347	
Category	Theme	Sub-Theme	Coded Word Count
		Category Total	27685
		All Identity	14227
	Identity	Identity Only	4221
	Identity	Cultural Roles	7220
		Non-Visible Disability	2786
-		All Terminology	8244
arra		Terminology Only	1011
Cult	Terminology	Dis- ability	5061
Ŭ		Disability VS Chronic Illness	1610
		Disclosure VS Openness	562
		All Advocacy	5214
	Advogovy	Advocacy Only	1493
	Advocacy	Perception Skew	1937
		Inspiration	1784

Stage 6 – Content Weighting

This stage addresses how relevant the subtheme (e.g. identity or cultural roles) may be to its superordinate theme (e.g. cultural) by categorising these by the percentage of coded words (see Table 4 and Table 5) so as to provide provisional focus to how commonly the themes occur in the transcripts and whether they require more in-depth review via IPA. The content weight (low, medium, high) provide an axis for characterising the risk matrix in the proceeding stage.

Table 4 - Table of Weighting Categorisation

Level of Weighting	Description			
1. Low	Low participant weighting is allocated to issues which achieved below the 33 rd percentile of relative theme to category total.			
2. Medium	Medium participant weighting is allocated to issues which are within the 33 rd and 66 th percentile of relative theme to category total.			
3. High	High participant weighting is allocated to issues which exceed the 66 th percentile of relative theme to category total.			

Table 5 - Table of Themes with Salien	y and Content Weighting (Example)
---------------------------------------	---------------------------	----------

	Salient	Themes	
Category / Theme / Sub-Theme	Absolute	Relative	Content Weighting
Cultural	81.7%	32.8%	
All Identity	42.0%	51.4%	
Identity Only	12.5%	15.2%	46%
Cultural Roles	21.3%	26.1%	79%
Non-Visible Disability	8.2%	10.1%	31%
All Terminology	24.3%	29.8%	
Terminology Only	3.0%	3.7%	11%
Dis-ability	14.9%	18.3%	56%
Disability VS Chronic Illness	4.8%	5.8%	18%
Disclosure VS Openness	1.7%	2.0%	6%
All Advocacy	15.4%	18.8%	
Advocacy Only	4.4%	5.4%	16%
Perception Skew	5.7%	7.0%	21%
Inspiration	5.3%	6.4%	20%

Stage 7 – Severity Categorisation

Interpretive Phenomenological Analysis is used to populate the categorisation of consequences within the risk matrix through the in-depth familiarisation the researcher will have with the data set. An example for characterising the severity outcomes are outlined in Table 6. A guideline, as adapted by Derrico et al, (2011), of characterisation of consequences are as follows: high priority outcomes can be topics that necessitate urgent resolution which may severely impede function and/or satisfaction of using a system; medium ranked severity might be concerns that are obvious but do not require immediate change; lower priority may indicate issues with simple or superficial solutions with negligible or limited risk. Input from System Matter Experts (SMEs) can be utilised to impart additional knowledge to appropriately weight risk factors and severity risks to actants and stakeholders.

Outcome Severity	Description
1. Minor	Negligible or Minor injury or adverse physical or mental health outcome.
2. Moderate	Moderate injury/ Moderate adverse physical or mental health outcome.
3. Serious	Serious injury/ Serious adverse physical or mental health outcome.
4. Major	Major Injury/Major adverse health physical (e.g. possibility of permanent function loss) or mental health outcome (e.g. major depressive episode).
5. Catastrophic	Possibility of death or permanent loss of function (motor, sensory, cognitive).

Table 6 - Table of Consequence Severity Outcomes (Adapted from Derrico, et al., 2011)

The traditional risk matrix (see Table 7) populated using the categorisation outlined in stage 6 and 7 can be used to apply priority weighting to the themes observed in the data set (see Table 8)

Table 7 - Matrix to Calculate Level of Priority

			Consequence				
		Minor	Moderate	Serious	Major	Catastrophic	
			1	2	3	4	5
	Low	1	2	3	4	5	6
Weighting	Medium	2	3	4	5	6	7
	High	3	4	5	6	7	8

Table 8 - Table of Salient Themes with Content and Priority Weighting

	Salient Themes			
Category / Theme / Sub-Theme	Absolute	Relative	Content Weighting	Priority
Cultural	81.7%	32.8%		
All Identity	42.0%	51.4%		
Identity Only	12.5%	15.2%	46%	High
Cultural Roles	21.3%	26.1%	79%	High
Non-Visible Disability	8.2%	10.1%	31%	Medium
All Terminology	24.3%	29.8%		
Terminology Only	3.0%	3.7%	11%	High
Dis-ability	14.9%	18.3%	56%	Medium
Disability VS Chronic Illness	4.8%	5.8%	18%	Low
Disclosure VS Openness	1.7%	2.0%	6%	Low
All Advocacy	15.4%	18.8%		
Advocacy Only	4.4%	5.4%	16%	Medium
Perception Skew	5.7%	7.0%	21%	Low
Inspiration	5.3%	6.4%	20%	Low

Stage 8 – Priority Categorisation

With all themes given an appropriate priority level, this can help to identify the urgency with which issues should be addressed. These can populate a table or register of issues generated by the interviews which outline the specific area of concern where the issue or problem occurs. Further useful information includes the number of participants which it affects, other feedback or quotations to support and justify recommendations (for example, see Figure 1).

Main Category	Specific Category	Issue	Additional Information	%	Severity	Weighting	Priority
Educational	Negative Attitudes from Staff	Staff have discriminated or displayed inappropriate behaviour towards students with disabilities	Discriminatory behaviour has been encountered to those with learning difficulties and those who have require medical leave. Negative interactions have included condescending and patronising behaviour and being dismissive of support or personal equipment requirements	29.63%	4	2	6

Figure 1 - Example of Concerns Register

Discussion

A prominent aim of phenomenological research is to offset researcher bias, which is inherently difficult to achieve by the level of immersion and interpretation in qualitative research. This approach seeks to systematically examine narrative units through dissemination into smaller codified nodes of content and appropriately semantically quantify them (Gbrich, 2007). As system ergonomics explores the interplay between actors, their social environment and the relationships between them, there is benefit to looking at narrative inquiry through a system ergonomic lens and applying complimentary reporting techniques common in human factors assessments.

The quantification of qualitative data is often criticised as some believe it undermines the integrity of the source material. However, qualitative analysis is complex and diverse, and often has ill-defined guidelines. Therefore, this proposed approach seeks to represent a method of extrapolating data from large verbal protocol data sets that can be used across many system ergonomics applications, such as user experience trials, health and safety analysis and exploring relationships of inter- and intra- personnel in sociotechnical systems. The clarification of super- and subordinate themes further allows for transferability and generalisability through transparency of the analysis and thus increases validity of the methodological rigour. Whole and unanalysed transcripts are not viable for presenting data – the proposed approach outlines guidelines that presents data quantitatively without compromising the integrity of the narrative materials. This is through the use of IPA which helps the researcher stay close to the data and explore the unique characteristics of each participant in combination with HCA to identify patterns across data sets and visualise data in an organised framework.

References

 Braun, V., Clarke, V., & Terry, G. (2014). Thematic analysis. In P. Rohleder, & A. C. Lyons, *Qualitative Research in Clinical and Health Psychology* (pp. 95-114). Palgrave Macmillian.
Brinkmann, S. (2014). Interview. In T. Teo, *Encyclopedia of critical psychology* (pp. 1008-1010).

New York: Springer. doi:https://doi.org/10.1007/978-1-4614-5583-7_161

- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* Thousand Oaks, CA: Sage.
- Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks: Sage.
- Derrico, P., Ritrovato, M., Nocchi, F., Faggiano, F., Capussotto, C., Franchin, T., & De Vivo, L. (2011). Clinical Engineering. In G. Gargiulo, & A. McEwan, *Applied Biomedical Engineering* (pp. 169-196). London: IntechOpen. doi:10.5772/19763

Gbrich, C. (2007). Qualitative Data Analysis: An Introduction. London: Sage Publications.

- Glaser, B. G., & Strauss, A. L. (2017). *Discovery of grounded theory: Strategies for qualitative research*. Routledge.
- Glaser, B. S., & Strauss, A. (1967). The discovery of grounded theory. London: Sage.
- Howard-Payne, L. (2016). Glaser or Strauss? Considerations for selecting a grounded theory study. . South African Journal of Psychology, 46(1), 50-62.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford: Oxford University Press.
- Lewis, S. (2015). Qualitative inquiry and research design: Choosing among five approaches. *Health promotion practice*, *16*(4), 473-475.
- Leximancer Pty Ltd. (n.d.). Leximancer. Brisbane, Australia. Retrieved 2019, from https://info.leximancer.com/
- Pandit, N. R. (1996). The creation of theory: A recent application of the grounded theory method. . *The qualitative report, 2*(4), 1-15.
- QSR International Pty Ltd. (2012). NVivo qualitative data analysis Software. Version 10.
- Rasmussen, J. (1997). Risk management in a dynamic society: a modelling problem. *Safety science*, 27(2-3), 183-213.
- Rubin, H. J., & Rubin, I. S. (2011). *Qualitative interviewing: The art of hearing data*. Sage Publications.
- Sparkes, A. C., & Smith, B. (2013). *Qualitative research methods in sport, exercise and health: From process to product.* . Routledge.
- Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* . Sage Publications.