

# THEMATIC ANALYSIS

---

In this handout we summarise thematic analysis and describe two case studies.

## Introduction to Thematic Analysis

### The six stages of thematic analysis

Braun and Clarke [p.35, 1] describe the six stages of thematic analysis below:

Phase	Description of the process
1. Familiarising yourself with your data:	Transcribing data (if necessary), reading and re-reading the data, noting down initial ideas.
2. Generating initial codes:	Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3. Searching for themes:	Collating codes into potential themes, gathering all data relevant to each potential theme.
4. Reviewing themes:	Checking in the themes work in relation to the coded extracts (Level 1) and the entire data set (Level 2), generating a thematic ‘map’ of the analysis.
5. Defining and naming themes:	Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells; generating clear definitions and names for each theme.
6. Producing the report:	The final opportunity for analysis. Selection of vivid, compelling extract examples, final analysis of selected extracts, relating back of the analysis to the research question and literature, producing a scholarly report of the analysis.

# Criteria for Good Thematic Analysis

Braun and Clarke [p.36, 1] provide a 15-point checklist of criteria for good thematic analysis:

Process	No.	Criteria
Transcription	1	The data have been transcribed to an appropriate level of detail, and the transcripts have been checked against the tapes for 'accuracy'.
Coding	2	Each data item has been given equal attention in the coding process.
	3	Themes have not been generated from a few vivid examples (an anecdotal approach), but instead the coding process has been thorough, inclusive and comprehensive.
	4	All relevant extracts for all each theme have been collated.
	5	Themes have been checked against each other and back to the original data set.
	6	Themes are internally coherent, consistent, and distinctive.
Analysis	7	Data have been analysed – interpreted, made sense of – rather than just paraphrased or described.
	8	Analysis and data match each other – the extracts illustrate the analytic claims.
	9	Analysis tells a convincing and well-organized story about the data and topic.
	10	A good balance between analytic narrative and illustrative extracts is provided.
Overall	11	Enough time has been allocated to complete all phases of the analysis adequately, without rushing a phase or giving it a once-over-lightly.
Written report	12	The assumptions about, and specific approach to, thematic analysis are clearly explicated.
	13	There is a good fit between what you claim you do, and what you show you have done – ie, described method and reported analysis are consistent.
	14	The language and concepts used in the report are consistent with the epistemological position of the analysis.
	15	The researcher is positioned as <i>active</i> in the research process; themes do not just 'emerge'.

# Case Study 1

## Stage 1: Data

The goal was to investigate how research mathematicians produce new mathematics in a collaborative online setting.

The data was Question 1 of MPM 2009 and Question 2 of the 2011 in the International Math Olympiad. The relevant websites are the research thread for the problem solving process, a discussion thread for meta-discussion about the project, and a wiki page for a summary of the problem and discussion. Q2 2011 was solved over a period of 74 minutes by 27 participants through 174 comments on 27 comment threads. We performed a thematic analysis to develop a typology of comments, using software "dedoose" to create hierarchies of tags. We tagged 559 excerpts.

## Stage 2: Codes

The screenshot shows the dedoose software interface. At the top, there's a navigation bar with icons for Home, Codes, Media, Excerpts, Descriptors, Analyze, Memos, Training, Security, Data Set, Back, and Projects. The 'Codes' icon is highlighted. On the left, a sidebar titled 'Selection Info' lists two entries: 'mpm1-2009 (34840-34999)' and 'meta-level comment abo...'. Below this is another entry: 'mpm1-2009 (34864-34868)' and 'emotion or value words'. The main area is titled 'Document : mpm1-2009' and shows a text excerpt from a conversation. The text discusses a mathematical proof involving steps and obstacles. Annotations are present in the text, such as 'gowers' and 'Re 60. Another small case.' A large bracket highlights a section of the text. The bottom right of the main area shows a 'Codes' sidebar with a hierarchical list of categories: comment type, clarification, concept, conjecture, errors, example, explanation, extension to the problem, goals, and justification. The 'comment type' category is expanded, showing its sub-categories. At the bottom of the interface, there are buttons for 'Font Size' (with a dropdown menu), 'Prev Excerpt' and 'Next Excerpt', and a 'Selection: (34866-34866) Create Excerpt' button.

## Stage 3: Themes

### A typology of comments

#### Conjectures:

**One can start with any point** (since every point of  $S$  should be pivot infinitely often), the direction of line that one starts with however matters!

**Perhaps even the line does not matter!** Is it possible to prove that any point and any line will do?

### A typology of comments

#### Concepts:

Since the points are in general position, **you could define "the wheel of  $p$ "**  $w(p)$  to be radial sequence of all the other points  $p' \neq p$  around  $p$ . Then, every transition from a point  $p$  to  $q$  will "set the windmill in a particular spot" in  $q$ . This device tries to clarify that the new point in a windmill sequence depends (only) on the two previous points of the sequence.

### A typology of comments

#### Proof:

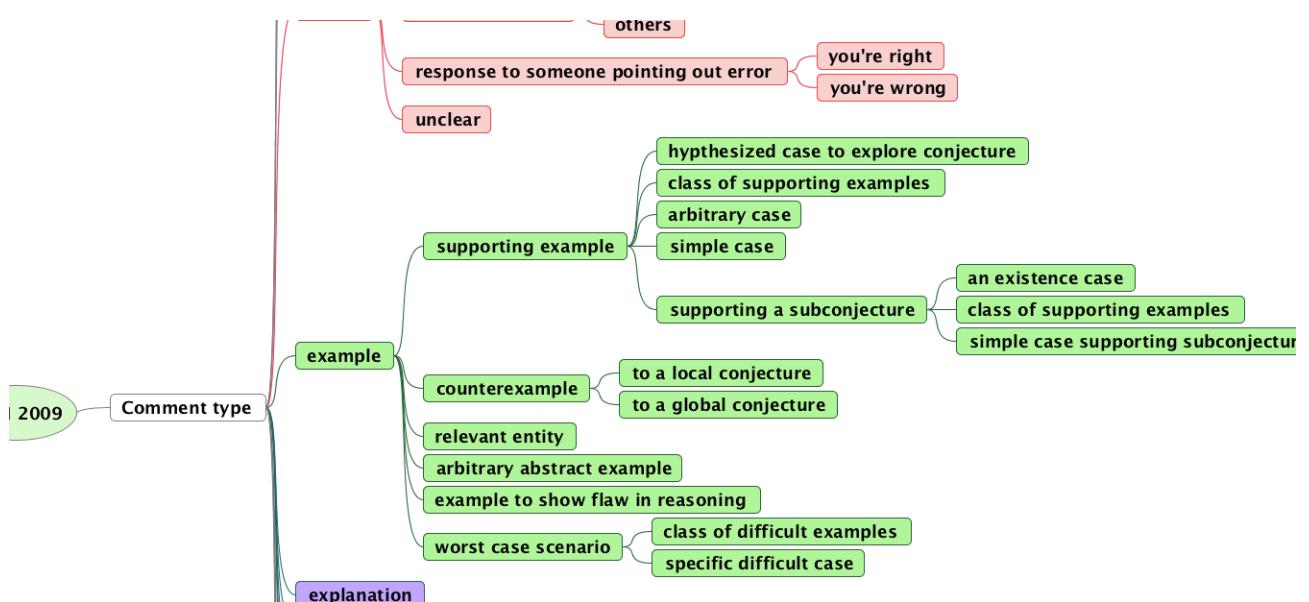
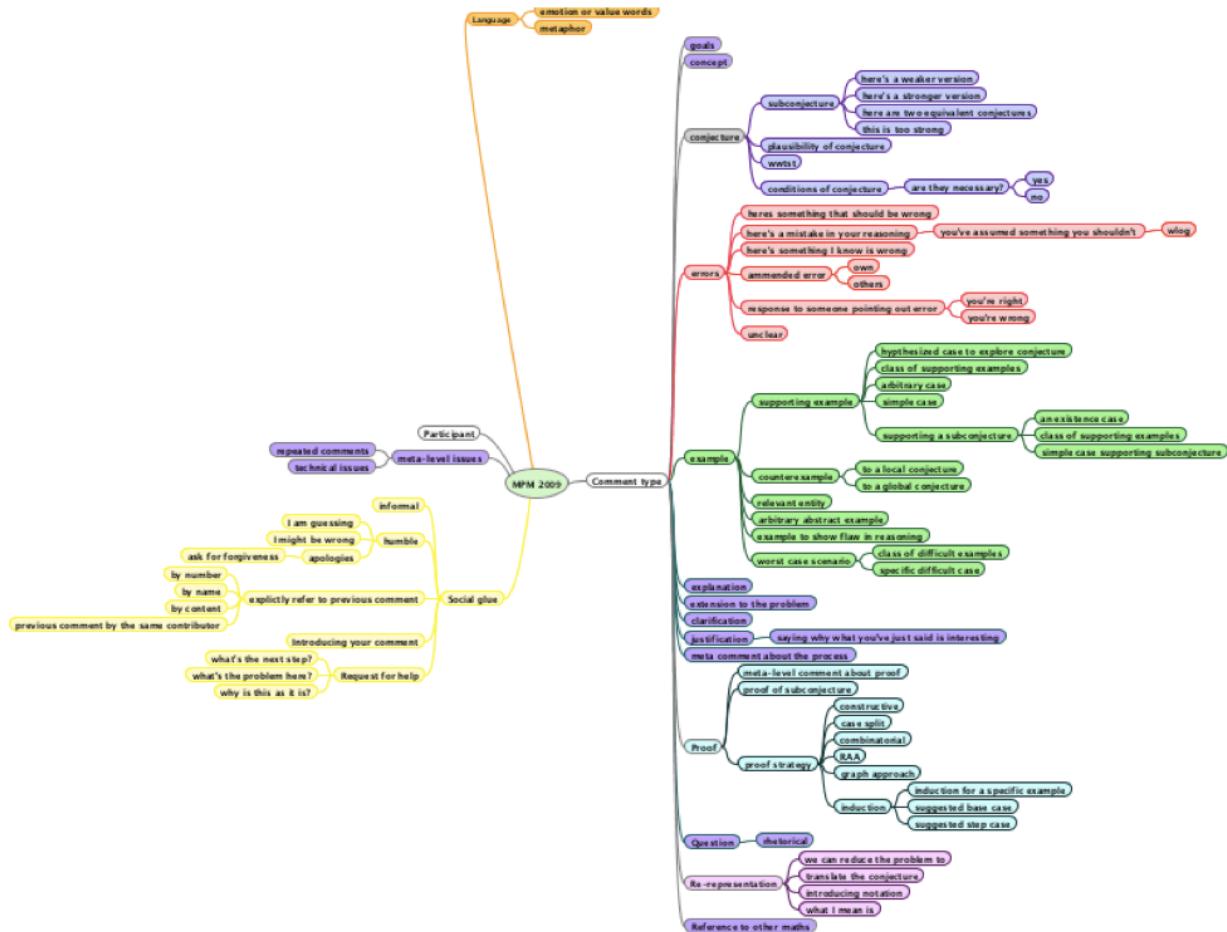
The first point and line  $P_0, l_0$  cannot be chosen so that  $P_0$  is on the boundary of the convex hull of  $S$  and  $l_0$  picks out an adjacent point on the convex hull. **Maybe the strategy should be to take out the convex hull of  $S$  from consideration; follow it up by induction on removing successive convex hulls.**

### A typology of comments

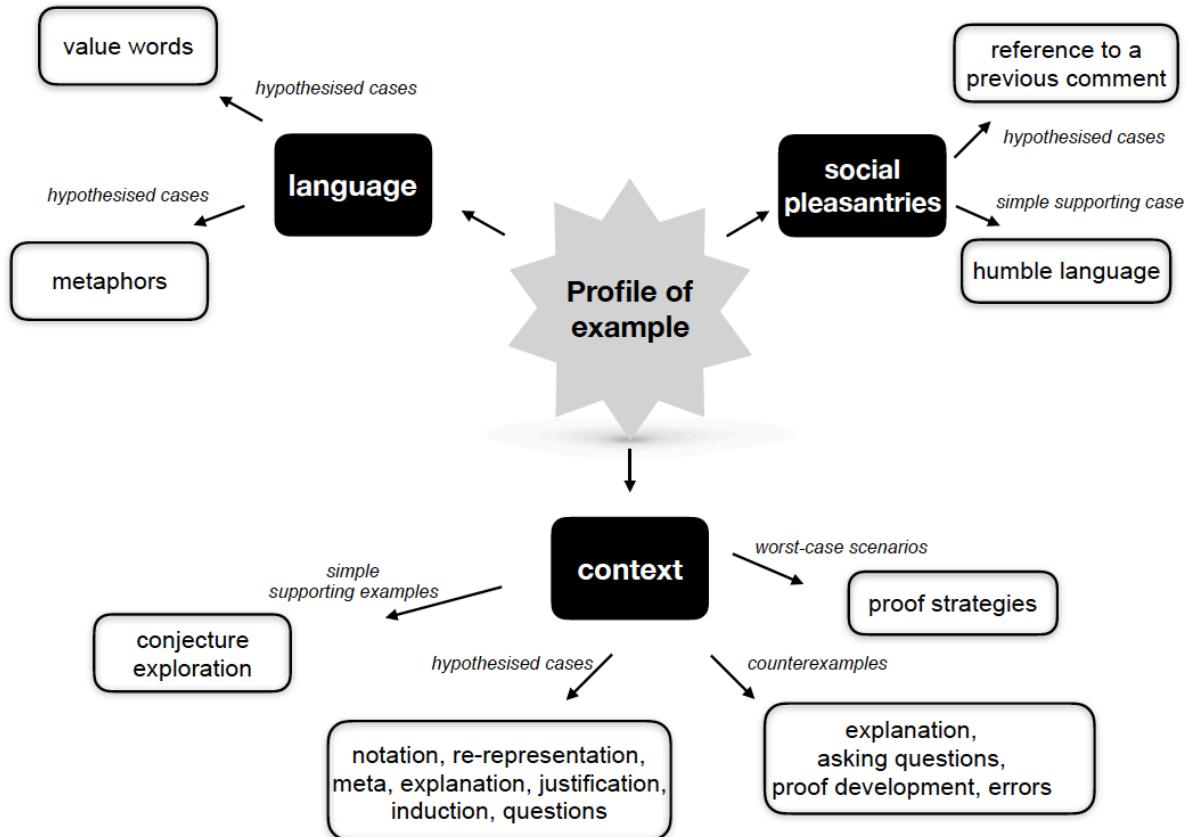
#### Examples:

If the points form a **convex polygon**, it is easy.

## Stage 4: Thematic Map



## Stage 5: Theory



## Stage 6: Report

Written up in [2], [3].

## Case Study 2

### Stage 1: Data

The goal was to examine how organisational factors affect the way its members respond to robots.

Data was collected over 18 months, via participant observations, fly-on-the-wall observations and interviews. The substantive population included hospital personnel, management, robot developers and patients.

### Stage 2: Codes

## OPEN CODING

Coding for concepts that are significant in the data as abstract representations of events, objects, relationships, interactions, etc.

Reliability analysis ensures objectivity of coding

Cohen's Kappa, >.70 acceptable

{abusing the robot}

I kicked it before [redacted] and I was told not to...  
[laughs]...when it first came. \*

\* Mutlu, B., & Forlizzi, J. (2008). Robots in Organizations: Workflow, Social, and Environmental Factors in Human-Robot Interaction. In Proceedings of HRI'08|—Winner of the best paper award.

### Stage 3: Themes

## AXIAL CODING

Concepts are categorized into explanations of arising phenomena (e.g., repeated events, actions, and interactions)

Negative treatments of the robot

- abusing the robot
- yelling at the robot
- impersonating the robot
- naming names

## Stage 4: Thematic Map

Show next page (→)

# SELECTIVE CODING

Integrate categories into a central paradigm—a “big picture” of the findings through building relationship across categories and contextualizing phenomena in data

Diagramming or tables could be used to build relational models

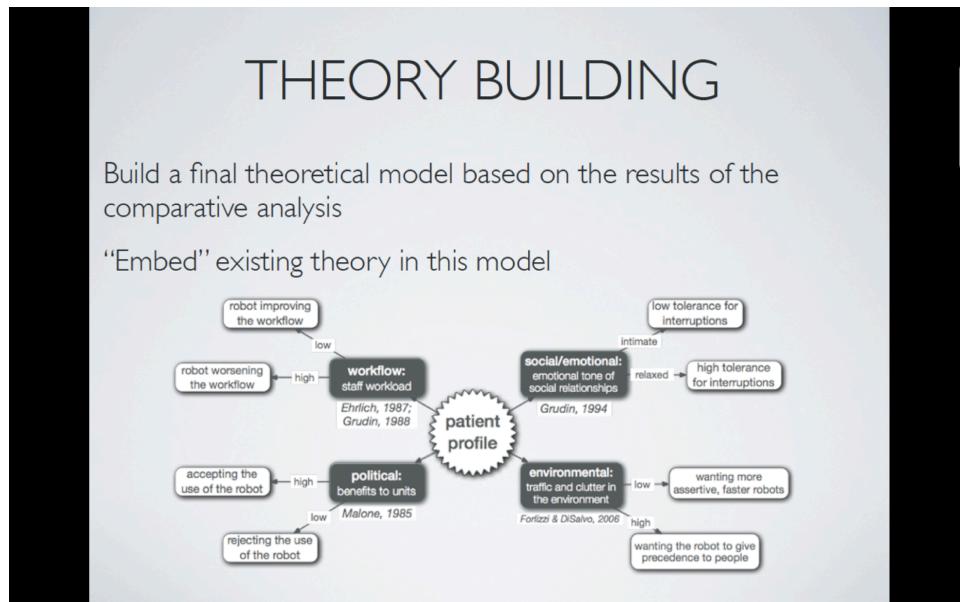
```
graph TD; A([Responses to the robot]) --> B([Conditions  
Heavy workload at units]); A --> C([Actions/Interactions  
Robot interrupting work]); A --> D([Consequences  
Negative treatments of the robot]);
```

# COMPARATIVE ANALYSIS

Compare the central phenomenon across several dimensions to understand how it is affected by social, physical, or organizational structures

```
graph TD; A([Perceptions of the robot]) --> B[negative  
Transcript from high-workload unit  
"[The robot] does tend to be annoying when [there are] 15 things are going on."]; A --> C[positive  
Transcript from low-workload unit  
"I think [the robot] is a delight I think it works fine, as it is."];
```

## Stage 5: Theory



## Stage 6: Report

Written up in [4].

## References

- [1] Braun, V and Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3 (2).
- [2] A. Pease and U. Martin. Seventy four minutes of mathematics: An analysis of the third mini-polymath project. In Proc AISB Symposium on Mathematical Practice and Cognition II pages 19-29, 2012.
- [3] A. Pease and U. Martin. “Human-like” example-use in mathematical research. Fourth Workshop on Bridging the Gap between Human and Automated Reasoning, in the joint Federated AI Meeting (FAIM) workshop program. (2018)
- [4] Mutlu, B. & Forlizzi, J. (2008). Robots in Organizations: Workflow, Social, and Environmental Factors in Human-Robot Interaction. In Proceedings of HRI’08