

Thematic Analysis and Artificial Intelligence: A Step-by-Step Process for Using ChatGPT in Thematic Analysis

Muhammad Naeem, Tracy Smith, Lorna Thomas

Research Background & Research Gap

Prior studies on AI in thematic analysis:

- De Paoli (2024a, 2024b): Used GenAI for initial coding and theme generation; found AI identified most main themes
- Lee et al. (2024): Used AI for interview transcription; only completed 3 of 6 steps
- Prescott et al. (2024): Compared human vs. AI coding quality; found human codes more reliable

Three research gaps identified:

- Did not complete all six steps of thematic analysis
- Did not provide AI with research contextual information
- Did not inform AI about methodological considerations at each step

Research Purpose & Contributions

Purpose:

- Develop ChatGPT prompts for each of the six steps of systematic thematic analysis

Two key contributions:

- Provides a complete prompt guide for all six steps; also serves as a training tool for novice researchers
- Demonstrates how to familiarize AI with research context (aims, theoretical framework, methodology, philosophical underpinnings) to reduce bias and improve transparency

Braun & Clarke (2006) vs. Naeem et al. (2023)

Braun & Clarke (2006)	Naeem et al. (2023) Systematic TA
1. Familiarization with data	1. Transcription, familiarization & quotation selection
2. Generating initial codes	2. Selection of keywords
3. Searching for themes	3. Coding
4. Reviewing themes	4. Theme development
5. Defining and naming themes	5. Conceptualization
6. Producing a report	6. Development of conceptual model

Key difference: Traditional TA takes a primarily **inductive** approach (themes emerge from commonalities in data); Systematic TA takes an **abductive** approach (themes are derived through interplay between data, research questions, and theoretical underpinnings).

The Six-Step AI-Assisted Process

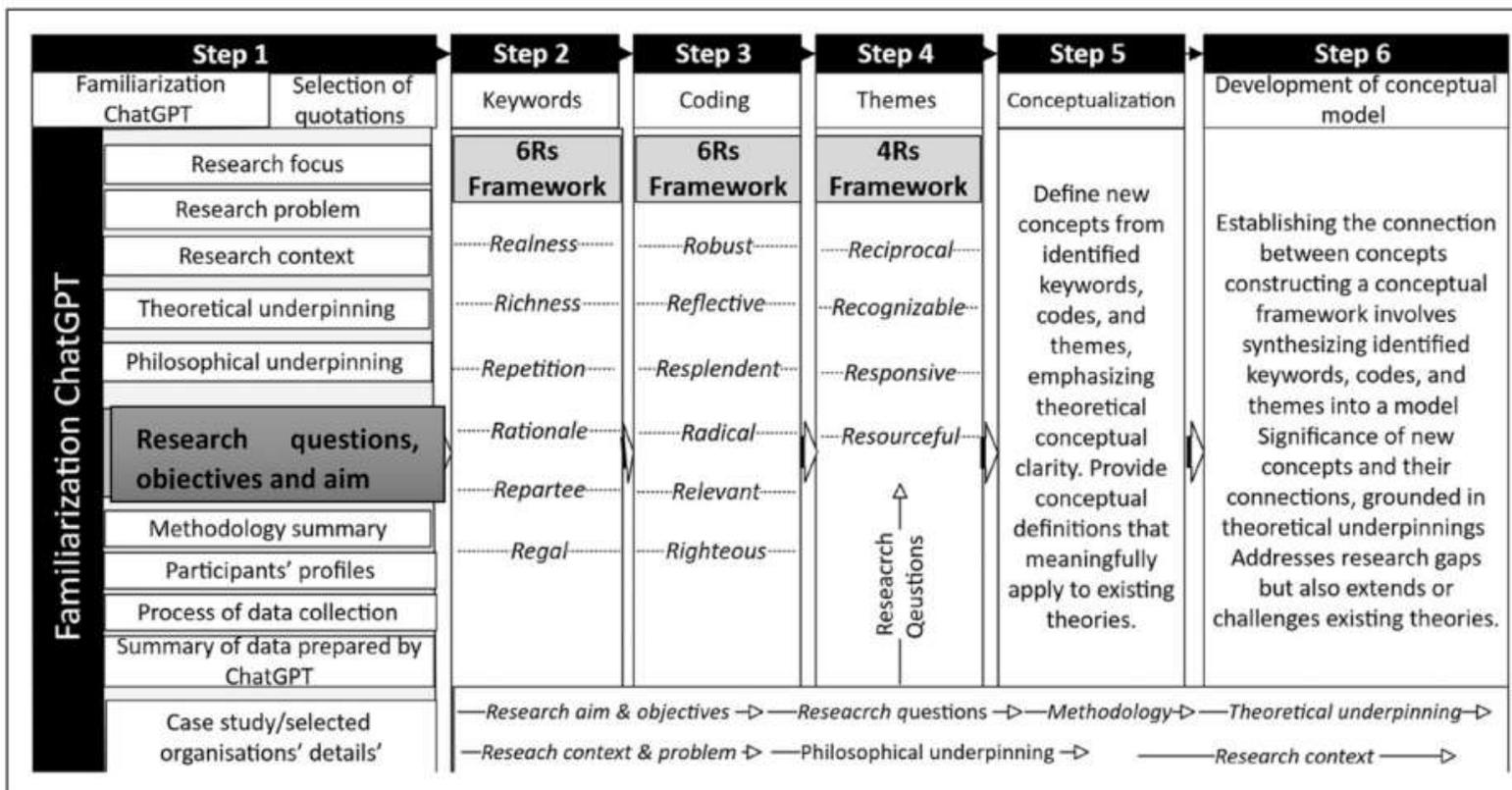


Figure 2. Using ChatGPT for Systematic Thematic Analysis: A Step-by-Step Process Using Artificial Intelligence in Qualitative Research.



Step 1: Familiarization & Quotation Selection

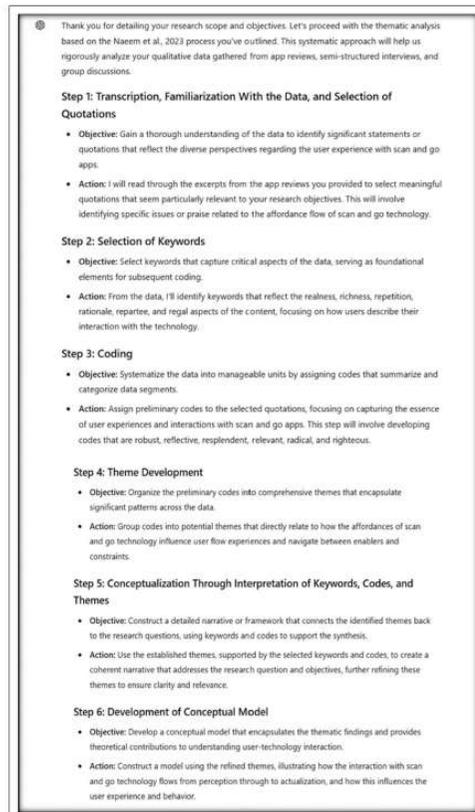


Figure 3. Confirmation of AI Familiarization With Systematic Thematic Analysis Framework (Source: ChatGPT (AI) Response).

Information provided to ChatGPT before analysis:

- Research aim, objectives & questions
- Theoretical framework (flow theory & affordance theory)
- Methodology (constructivist ethnographic approach)
- Philosophical underpinnings
- Data type & participant profiles
- Systematic thematic analysis process (Naem et al., 2023)

AI must understand the full research context before performing any analysis

Step 2: Keyword Selection — AI vs. Manual

- 6Rs Framework for Keywords: Realness, Richness, Repetition, Rationale, Repartee, Regal

ChatGPT Keywords	Manual Keywords (Naeem et al., 2024a)
Large quantity, specific, technical	Fewer, subjective, emotional
e.g., "Technical issues," "Digital coupons," "Navigation problems," "Screen brightness"	e.g., "Not my job," "Garbage app," "shopping life easy," "not for old and uneducated people"
Broader coverage across functional, experiential, and economic dimensions	Deeper capture of user attitudes and resistance

Step 3: Coding — AI vs. Manual

- 6Rs Framework for Coding: Robust, Reflective, Resplendent, Relevant, Radical, Righteous

ChatGPT: 16 Codes	Manual: 6 Codes
Specific, descriptive, function-oriented	Abstract, theoretical, psychology-oriented
e.g., "Technical Frustrations," "Budget-Conscious Shopping," "Navigation and Usability Issues," "Inventory Management Concerns"	1. Apathy 2. Excitement and risks 3. Gratification and capricious 4. Utility and hindrance 5. Absorption 6. Abandonment
Abductive reasoning (data-pattern driven)	Theory-driven (flow & affordance theory)

Step 4: Theme Development — AI vs. Manual

- 4Rs Framework for Themes: Reciprocal, Recognizable, Responsive, Resourceful

ChatGPT: 4 Themes	Manual: 3 Themes
1. Digital Shopping Experience Enhancement	1. Expectation and Perception
2. Operational and Technical Efficiency	2. Experience and Existence
3. Customer Engagement and Loyalty	3. Affordance Dichotomy
4. Accessibility and Inclusive Shopping	
Specific, action-oriented, descriptive	Abstract, theory-driven, narratively connected
Parallel categories (no clear progression)	Logical progression (before → during → tension)

Steps 5 & 6: Conceptualization & Conceptual Model

ChatGPT Concepts	Manual Concepts
8 concepts, largely overlapping with theme names	Affordances Flow Funnel:
e.g., "Digital Shopping Experience Enhancement," "Seamless Interaction Ecosystem," "Engaged Shopping Community"	Perceived Affordances → Actualization Affordances → Affordance Dichotomy
Descriptive, experience-based	Theoretical, process-oriented, with clear logical progression

- Step 6: Paper does not provide detailed AI vs. manual comparison for the conceptual model.
- Key implication: AI performs well at data processing level (keywords, codes) but struggles at the theoretical construction level (conceptualization, model building).

AI in Thematic Analysis: Benefits & Challenges

Benefits	Challenges
Reduces human bias and subjectivity	Algorithmic bias may replace human bias
Efficiently processes large datasets	"Black box" problem: internal process not transparent
Improves methodological consistency	Researcher influence through prompt design
Discovers patterns humans may miss	Reliability and accuracy require human verification

Authors' proposed solutions:

- Provide theoretical rationale for each prompt
- Use traditional quality control methods (member-checking, multiple coding rounds)
- Ensure thorough AI familiarization in Step 1

Critical Evaluation

Strengths:

- Addresses a genuine research gap with a clear rationale
- Highly practical: provides ready-to-use prompt templates with theoretical justification
- Sound comparison design: same data, same framework

Limitations:

- Single case study. Generalizability is questionable
- AI's weakness in theoretical construction (Steps 5 & 6) is underemphasized
- Author is both the method creator and case study author — potential conflict of interest
- No empirical evidence on actual time/efficiency gains