Recursive Abstraction Method for Analysing Qualitative Data

Recursive Abstraction is a method used for the analysis of qualitative data (Polkinghorne and Taylor, 2019). This qualitative data can be derived from data obtained from interviews (including group, semi-structured and unstructured interview formats), focus groups, and any other form of narrative based feedback, reports or accounts. As such, Recursive Abstraction is currently being applied in a wide range of discipline areas (Guzys et al., 2017; Ortiz et al., 2020; Rodriguez et al., 2020), and is particularly relevant to professionals and educators involved in Tourism Management and Marketing (Polkinghorne et al., 2020).

Often, in the case of Tourism Management and Marketing, qualitative data compiled for analysis is an expression of the views and opinions of those participating in the research, and so the data is inherently subjective. Whereas quantitative data is focussed on specifics, qualitative data is therefore more concerned with the identification of underlying patterns and trends, which need to be analysed in order for them to be revealed. Unlike other thematic analysis techniques, Recursive Abstraction operates as an iterative process in a series of repeatable steps. These steps include the identification of initial high-level themes, followed by the paraphrasing and coding of the data within these themes. As the data is collapsed, new emergent themes will often appear, which are then used to replace the original themes. The data is combined and collapsed repeatedly until a final set of Themes (high level groups or categories) and Codes (low level subject indicators) are revealed, which inform our understanding of the research phenomenon being investigated. The six steps of Recursive Abstraction are listed below (Polkinghorne and Arnold, 2014):

- **Step 1** – Highlight the data of interest within the text;
- **Step 2** – Extract the highlighted data and place it within a table or similar construct;
- **Step 3** – Paraphrase the data, i.e. use your own words to replace the data, whilst maintaining the overall meaning;
- **Step 4** – Group the paraphrased data relating to different topics together to form Themes. Everything within a Theme will be related in some capacity. For example, Themes relating to the feedback from guests staying at a hotel may include ‘arrival’, ‘accommodation’, ‘catering’, ‘activities’, and ‘departure’;
- **Step 5** – Replace the paraphrased data with Codes. Codes should encapsulate as much of the meaning in the paraphrased data as possible but using only a few words. The same Codes should be used whenever it is appropriate to do so. In this step it is important for the researcher to recognise that similar points may have been expressed in different ways. For example, “this is the best room ever”, “the room provides a
satisfactory level of accommodation” and “I like the room’s facilities” are all fundamentally a positive reflection on the room, and each could therefore be replaced by the same Code of “Great Room”.

*Steps 4 and 5 are repeated iteratively until the data analysis process is complete.*

- **Step 6** – The data that emerges from the final application of Step 5 are the findings of the research. These findings need to be reviewed in three ways:
  1. To check that no change in meaning has occurred to the data during the process of extraction, paraphrasing, grouping into Themes, and then into Coding;
  2. To identify patterns and trends in the data. For example, there might be a gender difference in the results, or perhaps the results for one hotel might be quite different from the results for another hotel, even though they have the same operational standards and processes;
  3. To develop recommendations which are based upon the findings. The output of research should always be to arrive at a better understanding of the phenomenon being investigated. Based upon this new understanding, change can be initiated.

As an example scenario, imagine a research project which explores the thoughts and views of holiday-makers using the local airport. The plan is to interview four different travellers at the airport as they each return from their summer vacations. From these interviews, the following comments have been identified (Step 1) and extracted (Step 2) for analysis:

**Younger Male Traveller:**
- “Booking the tickets is really easy and quick”
- “I like the really early flights as it means I arrive before lunchtime”
- “Baggage delays can ruin a holiday”

**Younger Female Traveller:**
- “Flight times were so early in the morning, but I don’t mind”
- “What cool destinations the company flies to now!”
- “Wow! how long did we have to wait to collect our bags!!!”

**Older Male Traveller:**
- “Limited parking space can be tricky”
- “Long security delays are a problem”
- “Booking the tickets is easy using the APP”
- “I now know to book early if I want to get the best flight times”

**Older Female Traveller:**
- “The company flies to some really nice destinations”
- “I wish the company would put on more flights”
- “I don’t like the long security queues”
- “There wasn’t enough parking when I needed it”
Once the data has been paraphrased (Step 3), grouped into Themes (Step 4), and then the paraphrased data replaced with appropriate Codes (Step 5), the final findings of the research will remain from which patterns and trends can be identified (Step 6). Example findings from Step 5 are illustrated in Table 1.

<table>
<thead>
<tr>
<th>Themes</th>
<th>Codes</th>
<th>Older Travellers</th>
<th>Younger Travellers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>Trip</td>
<td>Flight Times</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Destinations</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Arrival</td>
<td>Security Queues</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Baggage Delays</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planning</td>
<td>Parking Availability</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Booking Tickets</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Table 1 – Findings Summary of Themes, Codes and Patterns Identified in the Interview Data

Recommendations can now be deduced from the findings summary in Table 1. For example, older travellers are concerned about parking problems, and perhaps this could be addressed by a park and ride system, a larger parking area or advance parking reservation service. All travellers would value more flight time options, and maybe the range of flights could be expanded dependent upon demand. Female travellers valued the destination choices. Older travellers noted the long security queues, which could be addressed with more staff at busy times. Younger travellers were frustrated by the delays with baggage which could be addressed with more information and/or more staff. Male travellers appreciated the efficient booking system for tickets, and perhaps this can become a selling point in the future.

If a researcher wishes to identify ‘how many’ people agree with an aspect of the phenomenon being investigated, then a quantitative data method is the most appropriate choice. However, if a researcher wishes to understand ‘why’ people agree with an aspect of the phenomenon being investigated, then a qualitative data method is the most appropriate choice (Saunders et al., 2016). In this situation, Recursive Abstraction can play a pivotal role in the identification of patterns and trends hidden within the data.

Compared to other methods for qualitative data analysis, such as content analysis and thematic analysis, Recursive Abstraction is increasingly used within academic and professional studies (Thompson, 2018). The recursive nature of the process supports a more flexible approach to data analysis, enabling the process to respond dynamically to patterns and trends that emerge. Furthermore, the approach is simple to apply, and is effective in identifying patterns and trends hidden within the data, without the need to use specialist analysis software and tools.

References


