Content Analysis: Principles and Practices
**Steps in Content Analysis**

1. Identify data sources
2. Develop categories
3. Code data
4. Assess reliability
5. Analyze results

Documentation occurs throughout the process

---

**Step 1. Identify data sources**

- Identify appropriate data sources based on researchable questions
- Consider source and format of the information as well as its validity and reliability
- Decide upon unit of analysis - items to be coded (items from a survey, comment field from agency database, etc)
- Decide on selection of items to be analyzed – sample versus universe
**Step 2. Develop categories**

- Develop useful, valid categories based on researchable questions.
- Test a few items as a team (either independently or as a group) to identify items that may be ambiguous or still need further clarification.
- Code a subset of items using two independent analysts.
- Calculate preliminary inter-rater reliability.
- Refine categories, defining specifications about what should be included as well as what should not be included when there may be uncertainty to ensure reliability of coding.

**Step 3. Code data**

- Establish decision rule as to whether items should be coded to one category or multiple categories based on preliminary review of data and testing of coding.
- Decide how many analysts will code data, if two analysts will independently code the same data, or if a less rigorous method will be performed in that one analyst will code the data and another analyst will verify.
- Code data.
Step 4. Assess Reliability

- Upon completion of coding, assess inter-rater reliability
- For items in which there was disagreement, analysts should meet to discuss final disposition of coding by item
- For instances in which an agreement cannot be reached, a third person (e.g. AIC or AD) should review and arbitrate (blind reviews of previous coding recommended)

Step 5. Analyze results

- Based on coding scheme, review items by category – number of items, percentage of items coded to this category, themes that emerge
- Decide if categories should be “rolled up” to broader categories or themes
- Discuss categories that may be of special interest.
Class Exercise

Please turn to Appendix 2

Documentation of steps

Document steps taken throughout process, including:

• Selection of data sources
• Development of categories and subcategories
• Steps taken to ensure validity and reliability of both the data sources as well as categories used
• Steps taken to test and refine the categories
• Data analysis strategies
• Decisions made regarding the unit of analysis and number of items coded
Selected Qualitative Research Tools

- White Board
- Paper
- Sticky Notes
- Excel
- Access
- QPL
- NVivo

Please refer to Appendix 5

Reporting on Content Analysis Results

- Consider the level of specificity of the categories you report
- Keep results in the context of other findings
- Always report on the methodology used
- When using comments in a GAO report, pay attention to objectivity, fairness of the selection
Advantages of Content Analysis

- Systematic, transparent approach to using qualitative data
- Can make use of data that already exist
- Can use respondents' words to create reporting categories
- If documentary evidence is most appropriate, can use it in a systematic way

Disadvantages of Content Analysis

- Selected items may be incomplete or non-representative
- Bias may exist in creation of items (e.g., who answers open-ended survey items?)
- Inter-rater reliability may be difficult to achieve
- No "right/wrong" categories
## Appendix 1
### Steps and Considerations in Conducting and Reporting Content Analysis

<table>
<thead>
<tr>
<th>Steps in Content Analysis</th>
<th>1. Identify data sources based on researchable question(s) or sub-question(s) to be answered</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>What is the source of the information?</td>
</tr>
<tr>
<td></td>
<td>Will you use a sample or content analyze the population?</td>
</tr>
<tr>
<td></td>
<td>How will you obtain the information, and what kind of format will it be in? What is your unit of analysis? (For example, you might choose items from a survey, or comment fields from an agency database.) Once you've decided on your unit of analysis, decide on selection of items to be analyzed. Will you do a content analysis of a sample or the entire population?</td>
</tr>
<tr>
<td></td>
<td>2. Develop and refine categories</td>
</tr>
<tr>
<td></td>
<td>This step often takes a while, and is arguably the most important step in content analysis. It requires a similar amount of effort no matter which content analysis approach or tools you are using.</td>
</tr>
<tr>
<td></td>
<td>Category development requires iterations and multiple people are generally involved in identifying the categories.</td>
</tr>
<tr>
<td></td>
<td>Inter-rater reliability is important; you need to make sure the categories are the right ones, that is, useful given the researchable questions, and that more than one person would agree on the categories and their meanings. After a trial phase on a sample of the data to develop the coding categories (and before full coding begins), a work paper should be developed which contains explicit definitions of codes and any preliminary disagreement between coders.</td>
</tr>
<tr>
<td></td>
<td>An approach for doing this would be to:</td>
</tr>
<tr>
<td></td>
<td>• Test a few items as a team (either independently or as a group) to identify items that may be ambiguous or still need further clarification.</td>
</tr>
<tr>
<td></td>
<td>• Code a subset of items using two independent analysts.</td>
</tr>
<tr>
<td></td>
<td>• Calculate preliminary inter-rater reliability.</td>
</tr>
<tr>
<td></td>
<td>• Refine categories, defining specifications about what should be included as well what should not be included when there may be uncertainty to ensure reliability of coding</td>
</tr>
</tbody>
</table>
Pay attention to validity and reliability concerns both in creating/refining the categories and in placing the items into categories. This ensures the content analysis will be defensible and accurate, yielding evidence that will meet GAO's standards.

3. Code the data

The third step is to actually code the data. Each coder reads each item and places it into one or more category, depending on the coding scheme that is being used.

The decision on how many analysts will code data should be made with respect to time and resources available to the engagement team, risk level of the job, and if the analysis will provide sole support or corroborative support findings, conclusions, or recommendations.

4. In coding the data, a decision must be made as to whether two analysts will independently code the same data, or if a less rigorous method will be performed where one analyst will code the data and another analyst will verify each decision made by coder as to whether or not they agree with the code selected. Independent coding tests whether judgment is consistent between coders and ensures objectivity, and therefore requires limited additional corroboration. A less rigorous method likely needs to rely on more corroborating evidence.

5. Assess Reliability

The fourth step is to conduct reliability checks to ensure objective and accurate categorization process

Although attention to reliability was given during the first step as categories were developed and refined, overall reliability should be assessed to see how often the coders agreed (e.g. 80%).

Because of the attention in the beginning, the agreement between coders (which we refer to as inter-rater reliability) should be fairly high. The subsequent assessment of inter-rater reliability will determine the extent that the coders agreed on item placement into the categories, and will be used to identify disagreements that need to be resolved.
For items in which there was not agreement, the two analysts should meet to discuss their thinking about the decision-making process and why they each selected the category that they did. Often, this discussion leads to an agreement between the two analysts about which category is most appropriate.

For instances in which an agreement was not reached, a third person who did not participate in the initial coding (such as the AIC or AD), should review and arbitrate. It is recommended that the third person conduct a blind review – in other words, they will see which codes were assigned but they will not know which analyst assigned the corresponding codes. Blind reviews help to minimize subjectivity or biases the adjudicator may have regarding one of both of the analysts.

6. Summarize/analyze results

Based on your coding scheme, there are a various ways to summarize the results of a content analysis effort into information that will be useful in a GAO product. Although content analysis is much more than simply counting, you may find it useful to note the number and/or percentage of items in each category as you discuss themes that emerged. For example, noting number or percentage of items that have correlated responses (for example, do people who report being satisfied with their jobs also report higher raises?)

If there are a lot of categories, you may want to “roll up” the categories into broader categories, each capturing several subcategories – this is defensible, as long as you keep validity and reliability in mind, for example, by ensuring that more than one person would agree on the “roll up,” and making sure the items still fit into the larger categories. The level at which you report your findings should be consistent with your researchable questions.

Also, depending on your research and findings, you may want to focus on or highlight the contents of one or two categories that may be of special interest.

THROUGHOUT THE PROCESS - Document steps taken
All steps need to be documented for the workpapers, including how items were selected, how categories were developed and finalized, what steps were taken to ensure validity and reliability of both the categories themselves and the placement of items in them, and how the results were summarized.

**Considerations for Developing Categories**

- **How many categories?** (Should agree with goals of the job)
- **Do you want to “double code” items that may fall into more than one category, or make the categories mutually exclusive (and/or exhaustive)?**
- **Inductive versus deductive approach** (That is, do you know the categories ahead of time or not? For example, if you know you’re going to code opinion statements based on whether they are in favor of or opposed to a particular policy, you are likely to have two categories that you can identify up front (maybe three, if you want an ‘undecided’ or ‘mixed’ category. On the other hand, if you will be looking for what themes come up in open-ended comments, you probably know very little about possible categories ahead of time.)
- **What is the best unit of analysis** (e.g., the respondent or the idea)? (For example, if survey respondents give three or four answers, do you want to give each item its own category, or use the category that best captures everything the respondent said?)
- **Using a sub-sample of the responses to develop the categories** (A very good practice, particularly when there are a lot of items to be coded)
- **You may want to start with small categories and aggregate into larger ones, or vice-versa. Either is fine. Note that the categories you use to code items may be rolled up later for reporting purposes, if you want.**

**Considerations for Conducting and Documenting the Coding Process**

- **Keep track of decision rules to avoid ambiguity** – both to make sure all coders are clear about category definitions, and to make sure categories are transparent and defensible. Document and save all decision rules and discussions of category meanings for the work papers. This is the same GAO standard as if you are
Reporting on Content Analysis Results

- Pay attention to the level of specificity of the categories you report (Avoid rolling up some categories, but not others; this may bias results unintentionally.)

- Always report on the details of the methodology used, including how issues of external validity and inter-rater reliability could affect the results.

- When using comments in a GAO report, pay attention to objectivity, fairness of which comments are chosen. The context of the comment is important; make sure comments are fairly selected and that the reporting context is accurate.

doing computer runs on an agency database; all definitions and any changes have to be recorded.

- May need to revise categories during coding process – this is not unusual. However, avoid letting one or two weird items drive the whole category structure. Also note that if you change the meaning of a category you need to revisit all items previously coded in order to be sure the new category structure is still accurate for all items. (A good reason to work with a sample of items at the beginning, particularly if there are a lot to code).
## Resource Guide

### GAO Examples of Content Analysis

<table>
<thead>
<tr>
<th>Page</th>
<th>Report</th>
<th>Data type</th>
<th>Tool</th>
<th>Mission team</th>
</tr>
</thead>
<tbody>
<tr>
<td>A4-2</td>
<td>Auto Safety: NHTSA Has Options to Improve the Safety Defect Recall Process</td>
<td>Ten focus groups</td>
<td>Excel</td>
<td>PI</td>
</tr>
<tr>
<td>A4-3</td>
<td>DOD and VA Health Care: Federal Recovery Coordination Program Continues to Expand but Faces Significant Challenges</td>
<td>150 interviews</td>
<td>NVivo</td>
<td>HC</td>
</tr>
<tr>
<td>A4-4</td>
<td>Statewide Transportation Planning: Opportunities Exist to Transition to Performance-Based Planning and Federal Oversight</td>
<td>Open-ended survey questions</td>
<td>QPL</td>
<td>PI</td>
</tr>
<tr>
<td>A4-5</td>
<td>Environmental Health: High-Level Strategy and Leadership Needed to Continue Progress toward Protecting Children from Environmental Threats</td>
<td>Documents: meeting agendas, summaries, and letters</td>
<td>NVivo</td>
<td>NRE</td>
</tr>
<tr>
<td>A4-7</td>
<td>International Food Security: Insufficient Efforts by Host Governments and Donors Threaten Progress to Halve Hunger in Sub-Saharan Africa by 2015</td>
<td>Nine recorded panel discussions</td>
<td>Excel</td>
<td>IAT</td>
</tr>
</tbody>
</table>
Auto Safety: NHTSA Has Options to Improve the Safety Defect Recall Process, GAO-11-603, June 15, 2011

Auto manufacturers recalled more vehicles in 2010 than in any other year, according to the National Highway Traffic Safety Administration (NHTSA), the federal oversight authority for vehicle recalls. However, many recalled vehicles are never fixed, posing a risk to vehicle operators, other drivers, and pedestrians.

The Congress raised questions about the auto safety defect recall process, including the sufficiency of NHTSA's oversight authority and whether vehicle owners were effectively motivated to comply with recalls. In response, GAO reviewed laws and documents and interviewed NHTSA and stakeholders about the (1) extent of NHTSA's role in recalls and how its authority compared to that of selected federal and foreign agencies that oversee recalls, (2) benefits and challenges of the recall process for NHTSA and manufacturers, and (3) options for improving it. GAO also conducted focus groups with vehicle owners to better understand their perspectives.

GAO conducted ten focus group sessions with 89 vehicle owners at five geographically dispersed locations. These were structured small-group discussions designed to gain more in-depth information about specific issues than could easily be obtained by another method, such as a survey or individual interviews. The overall objective of the focus groups was to obtain the views, insights, and feelings of vehicle owners regarding their awareness of recalls, their understanding of defect notification letters, and their willingness to comply with defect notices. GAO hoped to meet this objective by including vehicle owners both with and without recent recall experience. All ten focus groups were recorded and transcribed.

The Content Analysis

The content analysis was conducted in two steps. In the first step, two analysts independently developed a code book and then worked together to resolve discrepancies. In the second step, each transcript was coded by an analyst, and a second analyst verified the codes. Coding discrepancies were resolved by the two analysts' agreeing on what the codes should be. Since the recall and nonrecall groups did not differ, totals for each question were compiled in a document that was used for the reported findings.
The focus groups were intended to generate in-depth information about the reasons for the focus group participants' attitudes toward specific topics and to offer insights into their concern about and support for an issue. Participants reported that (1) they preferred notification letters with certain elements and might be more likely to comply if the letters included VIN numbers and clarified the severity of the defect and (2) they were unfamiliar with NHTSA's primary means of communicating defect information to the public—its website. The ability to project the information produced by focus groups was limited (nonrandom selection of ten groups). Therefore, GAO used several different methods to corroborate and support its conclusions.

**DOD and VA Health Care: Federal Recovery Coordination Program Continues to Expand but Faces Significant Challenges,** GAO-11-250, Mar. 23, 2011

In 2007, following reports of poor outpatient case management at Walter Reed Army Medical Center, the Departments of Defense (DOD) and Veterans Affairs (VA) jointly developed the Federal Recovery Coordination Program (FRCP) to coordinate the clinical and nonclinical services that severely wounded, injured, and ill service members and veterans needed. This report examined (1) whether service members and veterans who needed FRCP services were being identified and enrolled in the program, (2) staffing challenges FRCP confronted, and (3) FRCP's challenges in coordinating care for enrollees. GAO reviewed FRCP's policies and procedures and conducted over 170 interviews of FRCP officials, Federal Recovery Coordinators (FRC), headquarters officials and staff of DOD and VA case management programs, and staff at medical facilities where FRCs were located.

**The Content Analysis**

For its content analysis, GAO used NVivo, qualitative data analysis software, to analyze more than 150 of 170 interviews with program officials and medical facility staff. The analysis helped GAO identify and quantify interviewees' responses on various topics. The program's coding capabilities made it possible to categorize interviewees' responses and provided a central place for reviewing and analyzing the documents.

GAO took a number of steps to ensure that the analysis was methodologically sound. First, potential categories were defined to organize the views of DOD and VA program officials and medical
facility staff by specific topics, including program eligibility criteria, the interviewees' interactions with the FRCs, overlap and duplication of activities, knowledge of the FRCs' role, and the challenges they faced. These categories were decided on by the themes GAO staff heard during interviews with the program officials and medical facility staff.

A preliminary intercoder reliability check was made to ensure the accuracy of the category definitions. To do this, two analysts coded a sample of 15 interviews into the categories. A methodologist compared the analyses to identify inconsistencies and, as a result, the categories that needed more specific definitions.

The same two analysts divided the final categories between them and coded the categories for all the interview documents. When they had completed the coding, each analyst reviewed all the codes the other had made and indicated whether he or she agreed or disagreed with them. Resolving their differences led to changes. Then they analyzed the Interviewees' responses by the defined categories. This analysis made it possible for the analysts to quantify the interviewees' responses within each category and support their findings on the third objective. That is, by analyzing interviews for common themes, the team was able to demonstrate that many officials viewed program eligibility criteria as unclear and that many officials had not made referrals to the program.

Statewide Transportation Planning: Opportunities Exist to Transition to Performance-Based Planning and Federal Oversight, GAO-11-77, Dec. 15, 2010

The states' transportation planning enables them to decide how to spend federal transportation funds—almost $46 billion in fiscal year 2009. Draft legislation to reauthorize federal surface transportation legislation would, among other things, have revised planning requirements to recognize the states' use of rural planning organizations (RPO) and would have required performance measurement.

Subsequently, GAO responded as follows to a request to examine (1) states' planning activities and RPOs' satisfaction that rural needs were considered, (2) states' planning challenges, (3) the U.S. Department of Transportation's (DOT) approach to overseeing statewide planning, and (4) the states' use of performance measurement and opportunities to base statewide planning more on performance.
GAO analyzed planning documents; surveyed departments of transportation in 50 states, Puerto Rico, and Washington, D.C., and 569 RPOs; interviewed officials in 6 states; and held an expert panel on performance-based planning.

To identify the extent to which state transportation departments were using performance measurement for planning, as well as opportunities to base statewide planning more on performance, GAO staff analyzed data collected through its state DOT survey and interviews with state transportation department officials. To gather information on the challenges the state departments faced in their statewide transportation planning, GAO relied primarily on data from the state survey, in which it asked state transportation department respondents to identify in open-ended responses the three most significant challenges to developing both the long-range statewide transportation plans and the state transportation improvement program (STIP).

The Content Analysis

GAO staff analyzed the content of the open-ended question responses by first grouping them into the 13 categories of challenges the state transportation departments had identified, including funding, stakeholder involvement, and staffing. Then GAO developed a codebook that defined each category, and two GAO analysts independently assigned codes to each response, resolving differences in their coding to reach consensus in a meeting. Then they removed duplicate responses—instances in which a state DOT reported the same challenge for the same plan more than once—to ensure that state departments reported only unique challenges. Finally, GAO staff analyzed the coded responses to determine how many state DOTs encountered each challenge in developing both the long-range statewide transportation plan and the STIP.

States commonly listed insufficient or uncertain funding for implementing transportation projects among the primary challenges to long- and short-range planning. They also reported that involving the public and addressing transportation data limitations were significant long-range planning challenges. Short-range planning challenges included meeting federal requirements to demonstrate the availability of sufficient project funding and to update the STIP to reflect changes.
Environmental Health: High-Level Strategy and Leadership Needed to Continue Progress toward Protecting Children from Environmental Threats. GAO-10-205, Jan. 28, 2010

Children face disproportionate health risks from environmental contaminants such as pollution in air, lead paint in homes, pesticide residues on food, and treatment-resistant microbes in drinking water. The contaminants contribute to asthma, cancer, neurodevelopmental disorders, and other diseases, and many of the nation’s 74 million children are exposed to them daily. In 2007, 66 percent of children lived in counties exceeding allowable levels for at least one of the six principal air pollutants that caused or aggravated asthma, contributing to medical costs of $3.2 billion per year, according to the Centers for Disease Control and Prevention.

In 1997, Executive Order 13045 had mandated that agencies place a high priority on children’s risks and required that policies, programs, activities, and standards address those risks. In response, the Environmental Protection Agency (EPA) created the Office of Children’s Health Protection and convened the Children’s Health Protection Advisory Committee.

The Content Analysis

GAO began by examining the extent to which EPA had institutionalized the protection of children’s health from the environment by (1) establishing agency priorities, strategies, and rules, including implementing Executive Order 13045; (2) using key offices and other child-focused resources, such as the Office of Children’s Health and the Advisory Committee; and (3) becoming involved in federal interagency efforts to protect children from current and emerging environmental threats.

GAO analysts then used NVivo, qualitative data analysis software, to analyze 35 Advisory Committee meeting agendas and related summaries derived from meetings held biannually or triannually between December 1997 and July 2009. They used the software also to analyze 74 Children’s Health Protection Advisory Committee letters sent to EPA and 53 EPA response letters issued between May 1998 and December 2008.
GAO subject matter and methodological experts developed a coding scheme for identifying (1) recommendations, defined as any and all statements made in Advisory Committee letters that advised, asked, requested, suggested, or urged EPA to take action and (2) EPA requests of the Advisory Committee, defined as formal or incidental requests for advice or input by EPA to its Advisory Committee. They identified recommendations in Advisory Committee letters to EPA. In some cases, a single sentence contained multiple recommendations. For example, the Advisory Committee wrote “EPA should show leadership in applying stringent mercury controls in our own coal-fired power plants and involve the U.S. in technology transfer to improve emissions in other parts of the world”; GAO coded this as two recommendations. EPA’s requests to the Advisory Committee were identified in meeting summaries, which represented the official and complete record of proceedings. Other requests—for example, individually from an EPA official to an individual Advisory Committee member—were not considered requests, since the entire Advisory Committee must be informed and must reach consensus on all matters, as specified in its charter.

GAO analysts then developed content analysis categories to characterize the range of issues the Advisory Committee recommended to EPA, based on a review of the Advisory Committee’s charter and an initial review of the letters. The analysts coded each recommendation into one or more of the following ten categories:

- budget and resources (financing, funding, or the need to change resource levels for a program or issue);
- education and public awareness (providing information to the public through different media outlets);
- organization and processes (how EPA is organized, including how it operates, the form or function of EPA management, and its internal processes and procedures);
- policies and priorities (advising EPA to amend, go forward with, or cease a particular policy or prioritization that could directly or indirectly affect children’s health);
- external partnership and interagency coordination (how EPA coordinated or collaborated with other agencies or entities);
- guidance (developing, updating, and using guidance documents and related information resources);
- regulations and standards (EPA regulations and its work setting or influencing EPA or government-wide standards);
- research (conducting, funding, using, or prioritizing research that would benefit children’s health);
• risk assessment (developing risk assessment protocols and selecting assumptions, risk factors, and margins of error); and
• tracking and indicators (tracking environmental pollutants, as well as monitoring such pollutants or observing human health outcomes over time).

The two analysts who conducted the content analysis discussed the discrepancies in their coding and reached agreement on them or resolved them through a third-analyst review. The final analysis produced an inventory of Advisory Committee recommendations and EPA requests of the Advisory Committee.

**International Food Security: Insufficient Efforts by Host Governments and Donors Threaten Progress to Halve Hunger in Sub-Saharan Africa by 2015, GAO-08-080, May 29, 2008**

In 1996, the United States and more than 180 world leaders pledged to cut 1990’s number of undernourished people globally in half by 2015. Subsequently, from analyses of U.S. and international agency documents, structured panel discussions with experts and practitioners, and fieldwork in four African countries, GAO was asked to identify (1) factors that contributed to persistent food insecurity in sub-Saharan Africa and (2) the extent to which host governments and donors, including the United States, were working toward cutting the region’s hunger in half by 2015.

For the first task, GAO used the United Nations (UN) Food and Agriculture Organization’s (FAO) estimates of the number of undernourished people and the prevalence of undernourishment—one of two progress indicators in the Millennium Development Goals target of halving hunger—to illustrate the lack of progress in reducing hunger in sub-Saharan Africa compared with other parts of the developing world. Further, GAO analyzed FAO’s data on input use, grain production, and grain planting areas to compare agricultural input use and productivity in sub-Saharan Africa and other parts of the world.

To summarize and organize meaningfully the many factors and interventions in global food security, GAO analysts created a framework from relevant literature (economic literature, studies, and papers issued by U.S. agencies, multilateral organizations, and bilateral donors). Nongovernmental organizations (NGO), government representatives in Washington, D.C., and officials from the four countries selected for fieldwork—Kenya and Tanzania in
East Africa and Mozambique and Zambia in southern Africa—reviewed the framework.

**The Content Analysis**

Choosing content analysis to complete the second task, GAO analysts conducted nine structured panel discussions in Kenya, Tanzania, Mozambique, and Zambia with about 80 participants representing more than 60 entities, including donors and NGOs. The same questions were posed to all nine panels, and their answers were recorded. The aim was to identify key recommendations for improving food security.

The analysts then coded recommendations and lessons according to the resulting factors. Regarding the recommendation to improve marketing, for example, GAO found that all nine panels mentioned it. GAO found further that not only were the recommendations and lessons, both positive and negative, mentioned at least ten or more times in at least six of the nine panels but that they were also consistent with the natural breaks in the data. GAO coded some recommendations and lessons learned according to a few additional topics that occurred with some frequency but that fell outside the scope of the framework.

Two analysts performed the initial coding independently and then met to reconcile differences. The views and perspectives of in-country NGOs, donors, and regional representatives could not be generalized beyond that population.