Chapter 5
Methodology and Data Sample

The present chapter consists of two parts. In the first part we shall examine what content analysis is and how relevant it is as a research methodology, and its application to journalistic and other media studies and some of the specific methodological issues related to content analysis. In the second part we shall discuss the specific methodology adopted for the present research.

5.1 Definitions

In this section we define some of the key terms and expressions that are pertinent to our study. These definitions are operational and set the limits and parameters within which we discuss the newspaper content. They serve as tools to explore or keys to unlock the concepts we would be discussing in the course of our analysis in the following chapters.

5.1.1 Content Analysis

Content analysis is a research technique that is based on measuring the amount of something in a representative sampling of some mass-mediated popular art form such as newspaper stories. It is a research method which involves counting communication phenomena and categorising them according to a taxonomy or
typology scheme. It is especially useful for archived data (texts, transcripts, audio and videotapes).

Content analysis is applied to a document consisting predominantly of texts and images, and which generally allows the researcher to analyse the material in a systematic manner enabling statistical comparisons to be drawn between subjects (Finn et al., 2000; Baker, 1994; Mehmetoglu, 2004). A textual analysis can be applied in several different ways, ranging from, for example, an interpretative analysis of literature by English literature students on a specific text, to a quantified form of analysis, known as content analysis (Veal, 2006).

Content analysis is the systematic quantitative analysis of communication content (Holsti 1969; Krippendorf 1980). It is used to determine the presence of specific terms or concepts in a text or set of texts and to infer meaning from such content in a given context. Analysis includes manifest content where the meaning is obvious, and the latent content where meanings are less apparent or indirect and may need interpretation.

According to Berger (1991) “Content analysis is a methodology by which the researcher seeks to determine the manifest content of written, spoken, or published communications by systematic, objective and quantitative analysis”. Content analysis has been defined as a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding (Berelson, 1952; GAO, 1996; Krippendorff, 1980; and Weber, 1990). Holsti (1969) offers a broad definition of content analysis as, "any technique for making inferences by objectively and systematically identifying specified characteristics of messages".
Under Holsti’s definition, the technique of content analysis is not restricted to the domain of textual analysis, but may be applied to other areas such as coding student drawings (Wheelock, Haney, & Bebell, 2000), or coding of actions observed in videotaped studies (Stigler, Gonzales, Kawanaka, Knoll, & Serrano, 1999). In order to allow for replication, however, the technique can only be applied to data that are durable in nature. Researchers are able to use content analysis to sift through large volumes of data with relative ease in a systematic fashion (GAO, 1996). It can be a useful technique to discover and describe the focus of individual, group, institutional, or social attention (Weber, 1990). It also allows inferences to be made, which can then be corroborated using other methods of data collection. Krippendorff (1980) notes that "much content analysis research is motivated by the search for techniques to infer from symbolic data what would be either too costly, no longer possible, or too obtrusive by the use of other techniques".

Texts include content of newspapers, magazines, books, articles, manuscripts, interviews, pictures and diagrams, transcript of conversations, advertised materials, theatre, and communicative language. In order to do content analysis, the text is usually coded, broken down into manageable categories and examined.

Cohen (1960) notes that there are three assumptions to attend to in using this method. First, the units of analysis must be independent. Second, the categories of the nominal scale must be independent, mutually exclusive, and exhaustive. Third, when using kappa is that the raters are operating independently. In other words, two raters should not be working together to come to a consensus about what rating they will give.
An important factor in any research is to recognise that a methodology is always employed in the service of a research question. As such, validation of the inferences made on the basis of data from one analytic approach demands the use of multiple sources of information. (Erlandson, Harris, Skipper, & Allen, 1993). Shapiro & Markoff (1997) assert that content analysis itself is only valid and meaningful to the extent that the results are related to other measures.

One advantage of this technique is that it is unobtrusive and not influenced by the presence of the researcher, as it is based on image or text and not behaviour (Babbie, 2001). In social researches in which individuals are surveyed or interviewed, the results can be influenced or the outcome can be affected by the very nature of the process of research. This does not occur in content analysis as the researcher is dealing with text and data that are more stable, static and not influenced. Another advantage is that the analysis can easily be repeated if any problems arise or if preliminary analysis was unsatisfactory (Babbie, 2001). Further, it is a safe form of data analysis because, if one makes mistakes in the process of doing content analysis it is possible to start all over again.

However, one major disadvantage of content analysis is that inferences made by the researcher can be subjective and it is possible that data can be used improperly to represent a cause rather than a reflection of socio-cultural phenomena. Researchers using content analysis further need to be conscious of anachronism affecting their research, especially when analysing ancient texts or data. If a text of the past is analysed with today’s tools and interpretative skills and knowledge, it is likely to affect objectivity. Space and time can affect the meaning of text. Language, which is
the substratum of much of the text under analysis, can undergo change with change in time and space. Hence, time, space and historical context associated with the data being analysed need to be carefully considered in order to avoid errors.

When used properly, content analysis is a powerful data reduction technique. Its major benefit comes from the fact that it is a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding and dealing with large volumes of data. The technique of content analysis extends far beyond simple word frequency counts. On account of the many limitations associated with word frequency counts, as we shall explain later, word frequency count has not been adopted in the current research.

Content analysis was used at first with newspapers towards the end of 19th century. It was done manually by measuring the amount of lines and space in newspapers. With the rise of common computing facilities like PCs, computer based methods became popular in content analysis. Answers to open ended questions, newspaper articles, political party manifestoes, medical records or systematic observations in experiments and a host of other data are today objects of textual analysis. Contents of communication be it from the mass media or from personal communications, available as machine readable texts, are taken as input and counted for word frequencies. Dictionary-based approaches generate a list of categories from the frequency list and control the distribution of categories over the texts. Methods in quantitative content analysis in this way transform observations of found categories into quantitative statistical analysis.
5.1.2 Prominence

The press considers prominence as one of the elements for an event or person to be newsworthy. The Oxford Advanced Learner's Dictionary (Hornby, 2000) defines prominence as the state of being important, well known or noticeable. Thus prominence is both an 'ascribed' as well as a 'perceived' quality. It is ascribed in as much as media and the public attribute importance to persons and events. It is perceived in as much as the audience views a person or event as prominent for a variety of reasons and on the basis of information they have obtained. Prominence in our study refers to the importance and attention the journalists, editors and the readers ascribe to news or other items in the newspaper. Prominence in our study is both ascribed as well as perceived.

The journalists make choices on what to report from the many events that happen or issues in any given day. The editors, in turn, make a series of decisions regarding what to publish and with what prominence, from among the number of items that come to the editorial desk. Importance of the person being covered in the news, the strangeness and rarity of the event, timeliness, element of conflict, the perceived interests of the readers, human interest are some of the principle factors that prompt both journalists and editors to cover news and information about certain persons or events. Competent journalists and editors are those who can gauge the proclivity of readers to read certain kind of items in the papers they subscribe to. Telling the story compellingly and powerfully to readers is vital to enhance reader interest. Undeniably, circulation and marketability greatly influence the choice of content, and the manner of presenting a particular event. All media, including
newspapers, create ‘heroes’ by attributing or giving prominence to individuals, often making them larger than life figures. Their subjects include political leaders, film stars, singers and musicians, sports personalities, to name a few.

Our study of prominence is based on measuring several factors, most of them latent, as they appear in the newspapers. We have examined the external, observable factors that are perceivable and measurable. We have also attempted to gauge non-latent factors like the perceived intention of the writer, editor or the reader. As we shall see in the following sections, we have used several methods to measure prominence. In our study prominence index is the sum of the values of the five prominence indicators each of which have a maximum possible value of 5. Hence for given news item, any of its particular attributes studied can have a maximum prominence index of 25.

We have also measured the prominence impact, which is a measure of the general impression regarding a given attribute of a news item created by a newspaper on the reader on account of the prominence factors. Hence we have termed it in our study as Prominence Impact. For a given attribute, prominence impact is obtained by multiplying the number of news items having that attribute by their average prominence.

5.1.3 Imbalance and Bias

Imbalance in our study implies an excessive attention given to a particular issue or perspective in the coverage of news and other items. Bias on the other hand, refers to decisions or acts based on prejudiced or pre-conceived notions and attitudes. According to Webster dictionary (2008) bias is an inclination of temperament or
outlook; a personal and sometimes unreasoned judgment, an instance of such prejudice, or deviation of the expected value of a statistical estimate from the quantity it estimates, systematic error introduced into sampling or testing by selecting or encouraging one outcome or answer over the other. It is not easy to distinguish or judge objectively bias since it can be present subtly. In our study we have examined bias and imbalance in news and editorial page items using several criteria. We have examined in great detail gender bias, rural-urban bias, religion bias as well as the imbalance in theme frequency.

5.1.4 Positive and Negative News

It is not always easy to divide newspaper content neatly into positive and negative news. We have tried to identify and distinguish the newspaper content as positive or negative on the basis of certain criteria we have described in the section on methodology adopted for the study. These criteria are guidelines that can help one to look at a particular report as related to an event that is positive or negative. We do not intend to go into the positive results of reporting a negative event or the effects these can have on the readers. Instead, we have categorised a particular news item as negative or positive on the basis of its content as empirically found in the newspapers under study. The public in general and readers in particular, identify news items as positive or negative on the basis of common perceptions and beliefs. An event is considered negative if it deals with violence, killing, kidnapping, cheating, corruption, environmental pollution and other criminal activities. Positive news includes development and service-oriented activities, promotion of peace, communal and religious harmony, environment, economic development, acts contributing to
enhance education, culture, health, human dignity, rights of people, and well-being of society.

5.2 Methodology Adopted in the Present Study

The present section discusses the methodology used for the content analysis, and procedures used to discuss and derive conclusions. The research explores what is covered in the press and how. Using a detailed set of criteria, the researcher has analysed the print version of news, editorials, articles and letters to the editor in the two newspapers chosen for research. The aim of the researcher is to study the various issues covered, the extent and nature of coverage, their impact, positive and negative characteristics and to examine the social responsibility of the press.

The study deals with measuring and analysis of content and aims at understanding newspaper content and its implications from the perspective of social responsibility. To do so, the researcher developed a methodology for data entry as well as analysis. The content was categorised as news and features, editorials, columns/articles and letters to the editor. A unit of research in this study is referred to by the term: item or story. Advertisements, obituaries, business and sports items were excluded from the purview of the study.

A multi-dimensional approach was adopted to measure various characteristics of stories. For every single story several factors have been measured. Values for each of the categories have been created after careful consideration. Taken together, these are expected to yield results that are objective and accurate.
The content analysis used in the present study is both quantitative and qualitative. It is quantitative, as the items have been counted for total numbers; they are measured for column weight; their location has been identified. Quantitative analysis of items include the total number of stories surveyed, their numbers, size, length, prominence in terms of location in the page, font size for headlines, frequency counts, space allotment, colour, photo size and page location.

It is qualitative in as much as the researcher looked at the specific meaning or significance of items on the basis of content, meaning and value. Qualitative analysis include examining whether a news item is positive or negative, whether it has been responsibly reported or not, its style, content quality, reliability, social responsibility, reporting ethics and so on. Compliance with or violations of ethical laws and regulations were examined so as to ascertain how responsible are the content of the newspapers. The potentially subjective nature of qualitative analysis was addressed in the research by the use of a detailed set of criteria related to several fields and by correlating them to determine qualitative factors. The researcher has adopted a combination of interpretive quantitative and qualitative method, which aims to investigate and interpret, within the context of the research subject, how the press depicts and mirrors society.

Qualitative features of items under study could be described as either latent or manifest. Study of ‘writing style’ for instance is considered as ‘latent.’ In content analysis normally categories should be mutually exclusive and not overlapping. At times it has not been easy to make fine distinctions. In such situations, what we could at best do was to affirm that a particular item indicates probability, trends, degrees of
certainty and so on rather than swear by absolute certainty. We have avoided the same story being coded under different theme categories. Items that could not be classified under specific category have been coded under ‘Other’ whenever considered necessary.

Aware of the various strengths and weaknesses of content analysis, the researcher has aimed to tackle the issues of subjectivity and ensure that inferences made about the data are carefully considered and maintained within the context of the subject matter. With regard to objectivity the comment made by Berlo (1960) is pertinent: "The scientist’s own values inherently are partial determinants of his work, the types of behaviors he chooses to study. In that sense, it is absurd to argue that scientific activity is value-free, or should be... The observer is part of any observation. That statement should lead the scientist to protect his observations as much as possible from his own biases, but it should not cause him to rule out his own experiences and introspective ideas from his conceptual framing of constructs and hypotheses."

The research involved a manual content analysis due to the non-availability of online versions of the complete newspaper texts. In addition, the researcher felt computer word searches inapplicable in the present research due to the varied categories studied and their interpretation to measure responsibility. Consequently, a computer word search, even at a preliminary stage, would not provide a detailed enough summary of the information provided by the newspapers, and hence it was felt that a physical reading of the print versions, though laborious and time consuming, would result in a fairer and more comprehensive analysis.
Even though there was a large quantity of data to be entered, since the newspapers under study were not available online or on soft copy, manual entry was adopted. The researcher with the help of a computer data entry operator manually scrutinised each of the 312 newspaper issues under study. The findings were entered in a computer programme designed for the purpose of our research using FoxPro. Though the process has entailed a lot of work and time, it has helped to ensure greater accuracy and reliability. Multiple coders were not involved since more than one coder would lead to the possibility of individual coders interpreting data and meanings differently, leading to coding inconsistencies. In order to avoid individual bias and coding inconsistency, all coding was done by the researcher himself.

We have measured many characteristics, such as:

1. Theme: The research deals with the theme of the item. It examines under the theme such questions as 'what is the story about' – through classification of news item categories with a view to find out the percentage of stories under various story types according to different themes (eg. How many numbers and percentage of news are related to politics; crime).

2. Location: The location of an item refers to the geographical focus which indicates whether a news item is international, national or regional, and again whether a news item is rural, urban or other. In location, the placement of the story in a particular page is also considered as it describes prominence.
3. Space: The amount of space allotted for each news category type in terms of percentages of pages is an important indicator to identify the prominence attached to various categories.

4. Size: Size refers to quantity. Quantity is measured on the basis of column numbers and space allotments and the size of the headlines.

5. Prominence: Prominence is studied from the point of view of column space and physical location. For instance, the first page items enjoy greater prominence in the estimation of both editors and readers.

6. Visual factors: Several visual factors in the newspaper add value to the content. These factors include presence of photos, and whether they are in colour or black and white; whether the item is in box or not, whether the headline is in bold font or not; typographical factors such as the font size of the headlines, whether the news is presented in screen (grey background or shading) or not. These attributes taken together or each of them separately enhance the prominence or importance of items the paper has covered.

7. Average number of stories per issue in each of the two papers has been identified. Comparative figures of stories between the two newspapers are also an important tool to gauge the content.

8. Gender Factors: We examined news items to see whether they are male, female or other category. Gender specifications are calculated in terms of numbers and percentages, and factors related to
prominence. Gender characteristic of an item is established on the basis of whether quotes, photos, names and themes in the story refer to male or female, or other. In examining information about gender we were able to cross check and correlate information to see other factors of prominence, namely headline weightage, column weightage, space allotment, presence of photos and so on. The data has helped us to examine if bias exists in news reporting on the basis of the number of stories, the nature of coverage, the manner of presentation and so on.

5.3 Research Design

The research project has been designed to undertake content analysis of news, editorials, articles and letters to the editor in *The Assam Tribune* and *The Shillong Times* spread over a 3-year period. The purpose of the research is to examine whether newspapers are socially responsible or not. In this section we wish to define the terms and categories used for our research. The study examines, analyses, evaluates content, presenting evidence to support conclusions related to the social responsibility of the press. The process of content analysis and the computer programme developed for this thesis can be applied to a variety of contexts and situations such as study of newspapers, books and other written literature, policy statements and documents, administrative records and so on. The researcher has developed a set of criteria which can be adapted and used according to need for choosing sample size, data
The study examines both quantitative as well as qualitative characteristics of the data contained in the two papers. A specific programme was developed to undertake content research to study published copies of the two newspapers. The data obtained thus is tested in the light of the hypotheses adopted for the research. Each individual item in the newspaper copies, which we have studied, has been examined from a variety of angles to examine the social responsibility of the press.

The coding system developed specifically for this study has been tested for consistency and accuracy. The computer programme enabled us to compute findings accurately and easily, as well as to retrieve data for re-examination or counter checking, thereby minimising the margin of error. According to the coding system we adopted, we assigned to each of the news items a specific identity enabling identification as well as prompt and accurate retrieval and counter checking.

The vast quantity of information recorded by means of the programme has enabled us to do a detailed content analysis and has offered scope for scientific analysis, interpretation of the data and deriving certain conclusions. This is expected to greatly reduce the margin of error associated with a study of this nature and magnitude.

5.4 Objectives

The press in India is closely associated with the freedom struggle, and many patriotic leaders like Gandhiji, Gokhale and others founded and edited newspapers.
The press in India played a crucial role in winning India independence. Since then the press has contributed significantly to preserve democratic values and foster national integration. Northeast India is geographically distant and isolated from the other parts of the country. Ethnically, the region is inhabited by a wide variety of tribes and ethnic groups having distinct identities, speaking diverse languages. The newspapers play a critical role in voicing the concerns and aspirations of the people of the region. They serve as the ‘Fourth Estate’ in democracy and other political processes like election, people’s participation and governance. The main objective of the present study is to assess the social responsibility of the press in Northeast India by analysing their contents. Therefore, the objectives of the present study are as follows:

1. To assess the social responsibility of the press in Northeast India by analysing newspaper content.
2. To measure the various kinds of bias in print media in Northeast India (gender, urban, local-national-international, religion).
3. To categorise the contents of newspapers in negative news (violence, conflict, killing, kidnapping) and positive news.
4. To study how prominently does the press cover social issues such as development, peace, education, health and environment.
5. To suggest ways to make the press more conscious of its social responsibility, and thereby fulfill its obligations to the public.

5.5 Hypotheses

While the study presumes that the newspapers under study have accepted the traditional roles and functions as their primary guiding principles, the study analyses and test certain hypotheses pertaining to the social responsibility of the press.
Keeping in view the above objectives of the study, the following hypothesis have been framed:

1. The print media in Northeast India is subject to bias (gender, urban-rural, and religious bias).
2. The press in Northeast India emphasises more on negative news (violence, conflict, killing, kidnapping) as compared to positive news.
3. Press is not socially responsible.
4. Social issues such as education, health, development, peace and environment are not prominently covered.

5.6 The Sample: Newspaper Identity

The research site of our study consists of the two dailies chosen are the oldest and most prominent papers of the two leading states of the region: Assam and Meghalaya. They are published from two of the most important state capitals of the region: Guwahati and Shillong, respectively. The period of study covers three years: 2003, 2004 and 2005.

We have chosen 52 issues per year of each of the newspapers, one for each week of the year. That gives us 156 issues for each of the two newspapers, giving a sample size of 312 issues of newspaper copies. The sample size of our study is 14.44 percent of the total number of issues of the two papers for the years 2003, 2004 and 2005. Each issue has been serially numbered from 1 to 312. Serial numbers 1 to 156 represents The Assam Tribune, 157 to 312 represents The Shillong Times. To ensure uniform distribution of the samples in the time period of the study one issue per week was chosen, thus obtaining 52 issues per year.
Further, in order to give equal weightage to each day of the week, each issue of a paper was chosen by lot from a series of numbers ranging from 1 to 7, representing the days of the week from Sunday to Saturday. We considered a cycle of 7 weeks. For each week in this 7-weekly cycle, one number was picked up at random from the series 1 to 7. After each selection, that number was eliminated for that 7-weekly cycle, to ensure that a newspaper of that day of the week was not chosen again for the same cycle. If the newspaper issue of a particular day in a week thus selected were not available due to a holiday on which the newspaper is not published or if the copy of a particular issue could not be obtained due to non-availability, the random selection was repeated till an available paper was chosen. This method ensured that each of the papers selected is of a different day of the week and that each day of the week was represented in the 7-weekly cycle. The process was repeated seven times in segments of seven weeks, to cover 49 weeks, and then repeated once more to cover the remaining three weeks of a year, thus selecting one issue for each of the 52 weeks.

This random and uniform distribution was adopted to ensure objectivity and impartiality and to avoid bias, to minimise distortion, to ensure accuracy and reliability. The process ensured that in any given year at least seven issues of a particular day’s paper would be studied.

5.7 Reasons for Choosing the Two Papers

The growth and development of the press in India is closely associated with the freedom struggle, and many patriotic leaders founded, edited or wrote for
newspapers. As has been noted in the first chapter, the press in India played a crucial role in winning independence for India. During the last six decades since India became independent, the press has contributed significantly to promote and preserve democratic values and foster national integration and secularism.

Northeast India is geographically distant and isolated from the other parts of the country. Ethnically, the region is inhabited by a wide variety of tribes and ethnic groups having distinct identities, speaking diverse languages. The newspapers play a critical role in voicing the concerns and aspirations of the people and fostering integration, development, harmony and peace. The researcher chose for the current research the two leading newspapers of Northeast India, namely, *The Assam Tribune*, from Guwahati, Assam, and *The Shillong Times* from Shillong, Meghalaya. These two newspapers are chosen, as they are the most important, among the oldest, most widely circulated papers of the region, and representative newspapers of the region. They are published from two important state capitals of the region, Guwahati and Shillong.

The period of 2003-2005 was chosen in the light of the fact that it marks the beginning of the new century and millennium. The period is significant as it represents over 50 years after independence of India. Today the press is undergoing rapid and sweeping changes. Television and internet are posing serious challenge to the survival of newspapers. Circulations are declining every year all across the world, with a few exceptions in countries like India. In the face of these challenges, it is imperative that the newspapers re-examine their roles and functions. This period also marks the rapid growth and widespread use of other electronic and satellite media of
communication, especially television and internet in the region. The study is expected to help the newspapers to examine the impact of these media on the press and to adapt themselves to the changes. Hence, historically and technologically the period under study is considered critical to enable analysis of the constants and changes that affect the newspaper industry and its functioning.

The study is undertaken with a view to examine if the press in Northeast India is socially responsible, whether it is upholding the traditional role of being the 'Fourth Estate', fulfilling its social responsibility, its role as a watchdog of democracy and a medium that aims to promote values of unity, communal harmony, peace, national integration, development, democracy and the rights and duties of all citizens. The study also intends to make proposals on how the press should be socially more relevant, committed and responsible. The findings and proposals of the study are expected to have far reaching consequences on media in general and newspapers in particular, and enable the press to assess its functioning as well as plan its future strategy.

The readers consider the two newspapers as socially responsible. By and large they continue to enjoy good reputation, popularity and readers' esteem. The public has a lot of expectations of these two newspapers as providers of reliable, credible information. Both papers were founded by pioneers with great national vision and patriotic objectives. They undertook a lot of risk, braved innumerable difficulties and faced hardships to establish and keep the papers running. In the light of the historical background, it is expected that these papers maintain high standards of journalism and social commitment. The two papers enjoy considerable public esteem. They
regard them as important and reliable providers of news and information and expect them to be socially responsible. The findings of the study, it is hoped, have important implications for newspapers in the country in general. The study is also expected to help critically assess the social responsibility of the press, and make proposals to improve the functioning of the press in the country with special focus on Northeast India.

5.8 Limitations

Due to practical difficulty of examining physically vast quantities of data, the study is limited to news and features, editorials, articles, columns and letters. Excluded from the research are advertisements, listings and announcements, obituary columns, poems, children's columns, sports pages, business and economy pages, Sunday supplements. Columns in the editorial pages which are reprints of columns published many years ago, and are reproduced for historic interest are also excluded. However, news items and features, editorials related to business, sports or other themes have been studied whenever they occur in the general pages.

A story is said to have 'jumped' when it starts on the front page and is continued on another page, usually inside the pages with a tag such as 'see page 3, or continued on page 6. When news items 'jump' or continue to run from the first page to another inside page, it is assigned the value due to the first page item. Again, news pages are divided into upper half and lower half on the basis of the folding of the newspaper page in the middle. The upper half news is considered to be more prominent since readers tend to consider them more important than stories below the
fold or the lower half. The quantitative values are assigned on the basis of the total space allotted for the item.

### 5.9 Variables Used in the Research

The researcher examined the newspaper content pertaining to news category under 13 variables, and those pertaining to contents of editorial page items under 6 variables.

#### 5.9.1 Variables Used for News Items

The 13 variables used for study of news items are numbered serially under the following four broad categories:

- **A Identification**
- **B Prominence Index**
- **C Reporting Index**
- **D News Classification**

The variables used for News items are:

1. Newspaper Identity (1-312)
2. News Item Identity (PPCNN)
3. Column Weightage
4. News Location
5. Space Allotment
6. Headline Weightage
7. Photo Prominence
8. Content- Positive, Negative, Other
9. Content Quality- Responsible or not
10. Geographic Location – International-National- Regional

11. Location- Urban-Rural, Other

12. Gender Male, Female, Other

13. News type (36 theme categories)

We shall explain each of them.

A Identification

Under this section on Identification, the researcher has set out criterion followed in the present study for identifying each particular issue of the newspapers chosen for the study as well as the distinct identity of each news item being investigated. Taken together, the newspaper identity and news item identity provide a complete reference to issue of the paper and the specific story in the issue being investigated.

1 Newspaper Identity

Under the section Newspaper Identity, newspapers are identified with their names, year, month, date, day and week. Newspaper Identity refers to the serial numbers of the newspaper issues chosen for study numbered serially from 1 to 312. This is further divided into two categories according to the two newspapers as follows: The Assam Tribune 1-156; The Shillong Times 157-312. Newspaper Identity includes the date, day and week of the newspaper issue selected for study (For a complete listing of the newspapers with their identity, see Appendix)
2 News Item Identity

Besides the newspaper identity, each unit of item (news, editorials, columns/articles, letters) is assigned a code number. This enables one to distinguish each item from the other and make it possible to refer to any item in case of necessity, and to maintain the distinct identity of each item. The papers under study have a maximum of 1 to 16 pages. Hence, they are numbered by pages serially from 01 to 16. Thus, any reference to 01 indicates that the reference is to page 1. C stands for Column number. The papers being studied are broadsheets, having maximum of 8 columns, hence C is a single digit entry and represents column numbers 1 to 8.

Each news item in a paper has been given a unique identity in the format codenamed PPCNN, where PP stands for the page number and C represents the column number in which the news item appears. NN is a number of two digits, identifying sequentially from top to bottom, a news item in a column. This has been called News Item Identity. Thus for instance, if the news item identity is listed as 01603, it identifies the news item on page 01, in column 6, and is news item number 03 from the top. This method further enables to distinguish each particular story under analysis, and makes it easier to refer to the story for any further reference, rechecking or verification if needed. ‘News Item Identity’ includes all entries under the study including contents of editorial pages. However, a separate matrix was designed to enter data related to the editorial page as it is quite different from news pages.
It may be noted that the researcher has used terms like 'news', 'story', 'item' throughout to refer to news, features, editorials, articles/columns, letters to the editor which constitute the content of newspapers under investigation.

Whenever a photograph with photo and caption alone (without a detailed descriptive report) occurs it is treated as one of the news items like the rest and apply the same criterion as in the variables. For news with accompanying photos, a different coding system and allocation of values has been used as shall be described later. Photograph without caption, or without it being part of a news item occurs, it has been excluded from the analysis.

B Prominence Index

Prominence index indicates a number of variables, which help to establish the importance of a particular news item in a given paper. The variables studied are: how much space is allotted to a particular news item, how prominently it is located, how big is the headline, and whether the news item is accompanied by photographs. The researcher has tried to quantify these variables in the study by allotting values to each of these variables in order to examine 'Prominence Index'.

'Prominence Index', helps to establish whether a news item is given prominence or not. The researcher is indebted to John C. Pollock whose measuring score for data on newspaper coverage of US twin tower attack post-September 11, has been adapted to suit the present study. Pollock (2005) in his community structure approach study on the newspaper coverage of Islam post-September 11, uses the expression 'Prominence Score'. I have borrowed the expression 'Prominence' from Pollock, but have modified the expression and called it 'Prominence Index' in the
present study. The researcher has created variables to measure the prominence of a news item. Prominence Index is computed as the sum of the scores obtained from the following: 1. Column Weightage, 2. News Location 3. Space Allotment 4. Headline Weightage, and 5. Photograph. Each of these variables is measured on a scale of 1 to 5. Thus one can obtain a maximum score of 25 from which one gets a measure of prominence given to a particular news item by the paper. This is called ‘Prominence Index’. We shall examine the specific criterion followed for each of these categories that help to find the Prominence Index.

3 Column Weightage (on a scale of 1-5)

Number of columns allotted by the newspaper for a particular item is a measure of its prominence. The more the number of columns a news item gets, the more prominent it is. Accordingly the following scores have been given for the news items on the basis of column weightage. If a particular item is in a box or has a screen it has been allotted an additional point each as it enhances prominence. The numerical ratings for the column weightage on a scale of 1-5 are as follows:

Table 5.1 Numerical Ratings for Column Weightage

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 columns or more</td>
<td>4</td>
</tr>
<tr>
<td>4 &quot;</td>
<td>3</td>
</tr>
<tr>
<td>3 &quot;</td>
<td>2</td>
</tr>
<tr>
<td>1&amp;2 columns</td>
<td>1</td>
</tr>
<tr>
<td>News in box or screen</td>
<td>1 (additional point)</td>
</tr>
</tbody>
</table>
4 News Location (on a scale of 1-5)

The position of the news in a newspaper indicates prominence. Prominence in newspapers may be compared to ‘primetime’ on television. Prominence on the basis of news location is decided on the basis of the way the editors allot prominence to a particular news item by placing it in a prominent position in the page, and on the basis of how readers usually perceive importance given to the news on the basis of what they chose to read first. The position of the news invariably indicates prominence. While everyone agrees that the upper half of the first page is the most prominent space in a newspaper, opinion could vary regarding the comparative merits of values for the rest of the space allotments of news items.

In considering allotment of values to the rest of the pages, the researcher has been guided by assessment of reader preference. Readers normally want to get to the most important news items as fast as they can, and so normally read the upper half of the first page, then second half of the first page, followed by page 3, upper half, and then all other upper half pages, and then the lower half pages. This distribution of value is based on the researcher’s understanding and assessment of reading habits of people as well as the editors’ choice of space allotment for news of varying importance. As page 2 usually contains classified columns and advertisements in the papers under study, it is less likely to be the preferred choice of average readers before they read other pages. The last page, which is normally a prominent page and which should get prominence after the first page, has been omitted from our study as this page is usually devoted to sports. As mentioned earlier, sports pages have been left out of the study.
The numerical value for the news location on a scale of 1-5 is as follows:

Table 5.2 Numerical Ratings for News Location

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Page-1 upper-half</td>
<td>5</td>
</tr>
<tr>
<td>Page-1 lower-half</td>
<td>4</td>
</tr>
<tr>
<td>Page-3 upper-half</td>
<td>3</td>
</tr>
<tr>
<td>All other upper-half –pages</td>
<td>2</td>
</tr>
<tr>
<td>Lower-half pages</td>
<td>1</td>
</tr>
</tbody>
</table>

5 Space Allotment (on a scale of 1-5)

Quantity of space allotted for an item is an important indicator of prominence attributed to that particular item. The more space allotted for particular news, the more prominent it will be. Here the researcher has recorded the quantitative measure of space allotted for a particular news item to check the prominence of that particular item. Jump stories (stories continued onto inside pages, photographs included along with the story or additional photos in the inside pages) are also taken into consideration when measuring the quantity of space allotment.

Table 5.3 Space Allotment and Score

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>½ page or more</td>
<td>5</td>
</tr>
<tr>
<td>¼ page or more</td>
<td>4</td>
</tr>
<tr>
<td>1/8 page</td>
<td>3</td>
</tr>
<tr>
<td>1/16 page</td>
<td>2</td>
</tr>
<tr>
<td>All the rest (O to 1/16)</td>
<td>1</td>
</tr>
</tbody>
</table>

6 Headline Weightage (on a scale of 1-5)

News headlines in newspapers have at all times enjoyed prominence. The size of the headline indicates the importance of the news. Newspaper headlines are an
important factor in a newspaper copy. Readers want to get to know the news quickly and briefly by reading the headline. It must also have a high visibility and hence the size is an important factor. It is said that a good headline not only tells the story, but sells the copy. With declining newspaper readership being reported globally, a good headline is a decisive factor. The bigger the headline, the more likely that the paper will be sold and consumers will read. Often newspaper vendors shout out the headlines to motivate people to buy the paper and thus sell more copies. The expression ‘making headlines’ which has become an idiom in English language further underlines the importance of the ‘headline.’ Television news headlines and the news bar have increased the pressure on the newspapers to write more catchy and imaginative headlines for their stories. The bigger the size of the headline, the more prominent the news. One point is allotted for the headline in bold as it adds to prominence. Headline weightage is measured on a scale of 1-5.

**Table 5.4 Headline Weightage and Score**

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5 cm or more</td>
<td>4</td>
</tr>
<tr>
<td>1 cm or 1.5</td>
<td>3</td>
</tr>
<tr>
<td>.5 cm or 1</td>
<td>2</td>
</tr>
<tr>
<td>Below .5</td>
<td>1</td>
</tr>
<tr>
<td>Bold</td>
<td>1</td>
</tr>
</tbody>
</table>

7 **Photo (accompanying news) Prominence (on a scale of 1-5)**

Photographs, which accompany a news report, indicate prominence of the news item. Like space allotment, we measure the quantitative space taken up by the photo in terms of columns. A large photo will be published along with the text of the
news only when the editors feel that the event is important to warrant much attention, or the persons featured and the occasion are significant. Publication of one or more photos also indicates that the event or person is important. In considering the value of photos we have noted the size of the photo or the space it has occupied in terms of columns, the page where it is placed, whether the photo is in colour or black and white. The total value of different variables as indicated below is added up to determine the score for the photo prominence.

Table 5.5 Photo Prominence and Score

<table>
<thead>
<tr>
<th>Description</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photos of 3 or more column size</td>
<td>3</td>
</tr>
<tr>
<td>Photos of 2 or 1 column size</td>
<td>2</td>
</tr>
<tr>
<td>Photos of Colour Photos</td>
<td>1</td>
</tr>
<tr>
<td>Photos of Photo on front page</td>
<td>1</td>
</tr>
<tr>
<td>Photos of No photo</td>
<td>0</td>
</tr>
</tbody>
</table>

C Reporting Index

Though content analysis is considered by many scholars to be a research methodology suited for quantitative analysis, we have adopted certain steps to measure the quality of the newspaper content on the basis of whether the content is responsible or not. The qualitative nature of the news items are examined on the basis of whether and to what extent they adhere and practice norms and guidelines governing the journalistic profession. Such norms and guidelines are derived from legislations, recommendations of press organisations, governments, or other bodies that regulate the press, long standing traditions and conventions, public perception on
what is in good taste and what is considered responsible. On the basis of this we have examined if a particular item is responsible or not.

News items are also examined for their content value. We have classified them as Positive (P), Negative (N) or Other (O) on the basis of the nature of the news item reported. The guidelines to decide whether a particular item is Positive, Negative or Other is based on guidelines adopted for the study as indicated.

8 Content Quality

Having examined prominence index, primarily on the basis of quantitative data available in the issues of the papers, we now turn our attention to analyse the reporting quality with regard to news items. The study aims at examining the social responsibility of the press. The press is expected to fulfill certain social obligations. The public expects the press to provide credible and accurate information. Content quality is analysed in terms of degrees of responsibility which yield the reporting index. Newspaper items have been classified as positive, negative or neutral. Content analysis is primarily a tool for measuring quantitative data. Neuendorf excluded the possibility of qualitative analysis of data in content analysis. According to Readership Institute’s Newspaper Content Analysis (2001) in order to consider content analysis to be objective, the idea must be capable of objective measurement. It is obviously difficult to measure objectively the content of a news item to examine if it is ‘positive’, ‘negative’, or ‘other’. Though it is difficult to establish objectivity, it is nevertheless not impossible. While undertaking to examine the qualitative factors of news reports like content quality, positive-negative-neutral etc. one will not be able
to get as objective and accurate data as in the case of quantitative analysis. However, the researcher has adopted several verifiable criteria to measure responsibility. Coder subjectivity, bias, possibility of overlapping and exclusion are likely to be present. Every effort has been made to minimise these possible errors and to code qualitative data in order to obtain as accurate a picture as possible.

Content quality is measured on a scale of 0 to 5 on the basis of the responsibility of each and every news item in the papers under study. Criterion for considering the responsibility/irresponsibility of a news, editorial or other columns is based on traditionally accepted rules, guidelines and conventions of journalism, ethical considerations and legal provisions within which a newspaper and the journalistic profession functions. Constitutions of most countries provide explicitly or implicitly, special protection and privileges to newspapers and journalists in the belief that they are expected to render public service related functions. Therefore the press is expected to possess a high level of social responsibility. The public has great expectations of the press which include providing accurate, unbiased, objective news and views. Hence we have adopted a method to examine the responsibility quotient of the news item on the basis of whether a news item adheres to socially, legally and ethically accepted norms of fair journalism and responsibility.

Content Quality is measured on the basis of whether there are violations from the point of view of responsible journalism. Accordingly various degrees of responsibility/irresponsibility in the news items and content of editorial page are divided into five categories. The criterion for considering an item under one or the categories is listed below.
The extent of non-violation/violation is measured on a scale of 0 to 3. (0, 0.5, 1, 2, 3.) Each item in the newspaper under study which does not have any violation (0) as per the list below, will automatically be given a rating of 4 points and considered Responsible.

A score on the basis of whether there are violations from the point of view of responsible journalism or not may be described as follows:

1. Responsible
2. Somewhat Responsible
3. Irresponsible
4. Very Irresponsible
5. Absolutely Irresponsible

Criterion for considering a news item, editorial or other columns is based on traditionally accepted rules, guidelines and conventions of journalism, ethical considerations and legal requirements. The press is expected to have a high level of social responsibility and the public has great expectations of the press, which include providing accurate, unbiased, objective news and views.

For the purpose of this study the researcher has classified the extent of violation on a scale of 0 to 3. (0, 0.5, 1, 2, 3.) Each news or other item in the newspaper under study which does not have any violation as per the list below, will automatically be given a rating of 4 points and considered Responsible. As a news item under study is found to violate any of the qualities of responsible journalism, score points will be deducted. Accordingly the following content quality scores are ascribed.
1. Responsible: A news item having a score of 4 (0 violation)

2. Somewhat Responsible: A news item having a score of 3.5 (any violation under category 0.5)

3. Irresponsible: A news item having a score of 3 (any violation under category 1)

4. Very Irresponsible: A news item having a score of 2 (any violation under category 2)

5. Absolutely Irresponsible: A news item having a score of 1 (any violation under category 3)

**List Indicating Extent of Non-Violation/Violation**

0

No violation

0.5

Not reporting both sides of the story

Not identifying the source clearly or giving vague source

No clear or vague attribution

Accused party’s view not quoted

1

Sources, data unverified news

Editorialising news (opinion, comment as news)

Not objective news (i.e, public relations items, publicity without adequate information content, has personality cult, is cheap and irrelevant, celebrity-centred)
Partial and incomplete, not constructive

Has little benefit to the public as information

Not truthful/ or is false

Not in good taste

Not fair comment

Is harmful and violates moral and ethical values

Supports, justifies acts of violence, crime or terror

Publishing sensational photos: naked or mutilated bodies

Publishing obscene pictures: naked, semi-naked figures

Mention by name of victims of rape, HIV- AIDS, loathsome sicknesses

Making provocative statements, comments leading to violence, killing etc.

Inciting to violence, abetting crime, riots or communal disturbance.

Violations of the Limitations to Freedom of Speech and Expression under the Constitution of India, Art. 19(1) (a) under the following clause may be brought under the above according to the nature, extent and degree of violation.

1 Sovereignty and Integrity of India 2 Security of the State
3 Friendly relations with foreign States 4 Public order
5 Decency and morality 6 Contempt of Court
7 Defamation 8 Incitement to an offence
9 Content Value

Measuring a news item as positive or negative on a scale is problematic. However the researcher has followed certain criterion to identify whether a news item is: Positive (P) Negative (N) Other (O). All news items are classified under three categories on the basis of content value as Positive (P) Negative (N) Other (O). This is a qualitative method of content analysis. The criteria to consider an item as Positive (P) Negative (N) Other (O) is as follows:

Positive News (P)

Reporting of following categories of news will be classified under Positive news:

1. Life Saving/ Rescue Acts/ Relief, Rehabilitation Aid
2. Social Awareness
3. Communal Harmony/National Integration
4. Improvement of Civic Amenities/ Social Awareness
6. Peace Efforts, Peace Meetings, Peace Marches or Rallies
7. Literacy/Education
8. Positive, Affirmative Legislation
9. All other positive news

Negative News (N)

Reporting of following categories of news will be classified under Negative news:

1. Ethnic Riots, Conflicts
2. Abduction, Ambush, Kidnapping
3. Rapes, Sexual Assault
4. Bomb Blast, Arson, Looting
5. Bandhs, Protests, Strikes
6. Physical Assault, Wounding, Injuring
7. Floods, Earthquakes, other natural disasters
8. Verbal Abuse
9. All other negative news

Other (O)

There are several news items, which do not have clearly identifiable value in terms of classifying them as positive or negative. News or features that cannot be clearly distinguished as Positive or Negative is considered as other. Such items are included in the category of Other.

**D News Classification**

News classification includes geographic location to which the news or information pertains. All news items are categorized into: 1 International 2 National 3 Regional. *International* pertains to news outside India; *National* to news about India excluding Northeast; *Regional* includes news about Northeast excluding the rest of India. If a story has multiple locations, the one that is predominant in the news will be chosen. In determining the geographic location of news, the decision is made not merely on the basis of the dateline in the newspaper. Dateline indicates the place from which the reporter files the story; hence dateline does not necessarily indicate the location where event reported took place. For instance a story may have a New Delhi dateline, but the news may be related to Northeast India. In such cases, it will be classified under Regional, and not under National. Similarly a national event taking
place in the region (National Games) is considered as national. Hence, geographic location of news is dictated by the overall content of a story rather than names of places mentioned in the story.

10 Location (Geographic)

1 International 2 National 3 Regional

International pertains to news outside India; National pertains to news about India (excluding Northeast); Regional includes news about Northeast (excluding the rest of India)

For the purpose of easy data entry in the computer database the following codes were used.

Table 5.6 Location (Geographic) and Code

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>International</td>
</tr>
<tr>
<td>2</td>
<td>National</td>
</tr>
<tr>
<td>3</td>
<td>Regional</td>
</tr>
</tbody>
</table>

11 Location (Urban/Rural/Other)

The study also examined the location of news to understand whether news content is Rural or Urban. Any place equivalent to a district headquarters or above are considered urban; places below a district headquarters, are considered as rural. Though the Census of India considers factors like population, civic amenities etc. to classify a place as rural and urban, the researcher has not adopted this criterion as it would entail practical difficulties and become cumbersome to check each of these factors for each of our story locations. News is also classified as Urban, Rural and
Other on the basis of where the event takes place, and from where the report is filed. In instances when a story does not pertain to urban or rural, or cannot be clearly identified as belonging to either of them, it is considered ‘Other’. The three categories are: R= Rural U= Urban O=Other.

For the purpose of easy data entry in the computer database the following codes were used.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>Rural</td>
</tr>
<tr>
<td>U</td>
<td>Urban</td>
</tr>
<tr>
<td>O</td>
<td>Other</td>
</tr>
</tbody>
</table>

12 Gender

Newspapers are often blamed for gender imbalance and gender bias. One of the effective ways to study the treatment of gender is to count how many news items are about men, how many about women, and how many are other. Hence all the stories have been classified under one of the three categories. A consistent pattern has been followed in assigning a story under any one of the categories on the basis of content. Whenever the chief character of the story is female, or a woman presides over functions, or they are quoted in a significant way to affect the news which is reported, it is classified as Female. Whenever women’s photographs are there, or activities, programmes, seminars etc. are concerning women, or when the editorials deal with women related issues, we have classified them as Female. Correspondingly, when men are quoted, featured or primarily the theme is about men, it is classified as
Male. Stories, themes or photographs which have mixed characters (men and women) in the same report, it is classified as Other. M= Male  F= Female  O=Other.

For the purpose of easy data entry in the computer database the following codes were used.

**Table 5.8 Gender and Code**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>M</td>
</tr>
<tr>
<td>2</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>O</td>
</tr>
</tbody>
</table>

13 Theme

An item is classified under one of the many major themes categorised for the purpose of our research. All possible newspaper themes are categorised under a total of 36 broad themes, which are subjected to content analysis. The choice of the themes was based on the researcher's observation of various kinds of topics usually covered in the two newspapers. In order to standardise the criterion for deciding the content themes, certain boundaries for themes were set, which seem apparently overlapping. For instance, the theme 'politics' is differentiated from 'government' on the basis of whether a news item deals with party politics or administrative acts, policies and decisions of the government. A member of the government, say, a chief minister speaks at a political rally of his party, would be considered a news item under the theme of politics.

Content is an important factor for readership as well as reader preference. We have identified the themes corresponding to the news type. A separate list of items
with a serial number list was developed in order to enter frequency of various themes in the news items studied. In the course of data entry, the corresponding numbers which stands for the story has been entered. Most stories in a newspaper are complex and therefore involve more than one theme. The researcher had to decide which theme is most prominent in a particular story. At the time of coding the following questions were asked: what is the story about, what are its main points in terms of content, what is the central concept described in the story, in order to decide under which theme it should be classified.

In most instances, the story itself suggests its theme. In other cases, it is necessary to look for the news peg, the angle, the purpose or the main focus of the story which suggest the theme. The decision to include a story under one or the other themes is based not merely on reading the section titles, headlines, subtitles or highlights. When a story has multiple themes: Example: Puja being offered by Hindu religious priests to obtain victory in a cricket match, it could be classified under Sports or Religion. But in this particular instance, the religion angle is the more dominant one, which is the main reason why the news has found a place in the newspaper, and hence, it would be classified under Religion. When a particular news item has common denominators (e.g. socio-economic and political at the same time) the most prominent category on the basis of content is chosen. Since the coding of all items under research were done by the researcher himself it has been easier to follow coder consistency in classifying stories according to the theme categories. In the column only the numbers corresponding to the news type or theme have been entered.
To handle vast number of topics, the researcher has also grouped some topics of similar nature under one number. (For instance, Number 8 includes all news items pertaining to Accidents, Disasters, Floods and Earthquakes). The 36 theme classifications in the study as listed below have been later listed in the descending order with the highest number to lowest number of frequencies. We have also calculated the percentage of news items for each theme as against the total number of stories in order to establish the prominence given to a particular theme (See the chart with list of themes, frequency count and percentage). Further the 36 themes have been brought under four major categories using a taxonomy to further investigate the hypothesis adopted for the study.
Table 5.9 Story Themes

<table>
<thead>
<tr>
<th>S.No</th>
<th>Story Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Crime/Violence/Conflict/War/Corruption/Kidnapping</td>
</tr>
<tr>
<td>2</td>
<td>Politics</td>
</tr>
<tr>
<td>3</td>
<td>Administration/Government/Election/UN</td>
</tr>
<tr>
<td>4</td>
<td>Militancy/Terrorism</td>
</tr>
<tr>
<td>5</td>
<td>Education</td>
</tr>
<tr>
<td>6</td>
<td>Media/Books/Literature/Films</td>
</tr>
<tr>
<td>7</td>
<td>Health</td>
</tr>
<tr>
<td>8</td>
<td>Accidents/Disaster/Floods/Earthquakes</td>
</tr>
<tr>
<td>9</td>
<td>Celebrity/Stars/VIPs/Entertainment</td>
</tr>
<tr>
<td>10</td>
<td>Development/Aid/Water/Relief</td>
</tr>
<tr>
<td>11</td>
<td>Art/Culture/History/Language/Music/Cultural Festivals</td>
</tr>
<tr>
<td>12</td>
<td>Trade/Business/Industry/Tourism</td>
</tr>
<tr>
<td>13</td>
<td>Agitation/Dhama/Bandh/Strike</td>
</tr>
<tr>
<td>14</td>
<td>Bilateral Relation/Internal Security</td>
</tr>
<tr>
<td>15</td>
<td>Environment/Nature/Forest/Ecology/Climate/Weather/Wildlife</td>
</tr>
<tr>
<td>16</td>
<td>Law/Justice/Human Rights</td>
</tr>
<tr>
<td>17</td>
<td>Religion</td>
</tr>
<tr>
<td>18</td>
<td>Police/Armed Forces/Military/Defense</td>
</tr>
<tr>
<td>19</td>
<td>Transport/Communication</td>
</tr>
<tr>
<td>20</td>
<td>Science/Technology/Space/Nuclear Issues</td>
</tr>
<tr>
<td>21</td>
<td>Economics/Finance/Budget/Banking</td>
</tr>
<tr>
<td>22</td>
<td>Peace/Harmony</td>
</tr>
<tr>
<td>23</td>
<td>Games/Sports/Adventure</td>
</tr>
<tr>
<td>24</td>
<td>Agriculture/Food/Pisciculture/Animal Husbandry</td>
</tr>
<tr>
<td>25</td>
<td>Law and Order</td>
</tr>
<tr>
<td>26</td>
<td>Fashion/Sex/Beauty Contests</td>
</tr>
<tr>
<td>27</td>
<td>Migration/Influx/Refugees</td>
</tr>
<tr>
<td>28</td>
<td>Employment/Career</td>
</tr>
<tr>
<td>29</td>
<td>Communalism</td>
</tr>
<tr>
<td>30</td>
<td>Women</td>
</tr>
<tr>
<td>31</td>
<td>Social Issues/Ethics</td>
</tr>
<tr>
<td>32</td>
<td>Death/Killing/Murder</td>
</tr>
<tr>
<td>33</td>
<td>Children</td>
</tr>
<tr>
<td>34</td>
<td>Youth</td>
</tr>
<tr>
<td>35</td>
<td>Nation/Patriotism</td>
</tr>
<tr>
<td>36</td>
<td>Population/Ethnicity</td>
</tr>
</tbody>
</table>

5.9.2 Variables Used for Editorial Page Items

The contents of editorial page items are studied under six variables. As the content and format of the opinion and editorial page is different from the rest of the
news pages, a different approach to study its content is followed. A separate matrix has been followed in order to enter data about the editorial pages and to analyse them. The three commonly found items in the editorial page are: 1 Editorial, 2 Columns/Articles, 3 Letters to the Editor. Each of them is assigned a number in serial order. Hence, the numeral 1 represents editorials, 2 articles, and 3 letters to the editor. Individual items within each of these would be given an extension marked with the number plus a point and serial numbers. Editorials in general are assigned Number 1. Hence, 1.1 would indicate the first editorial; 1.2 would indicate a second editorial, if there is a third editorial it would be 1.3, and so on. Number 2 indicates articles. Counting for articles, signed or syndicated columns, the numbering is from left to right. Accordingly, 2.1 refers to the first article, 2.2 the second, 2.3 the third and so on. Number 3 represents letters to the editor. The letters, counting from left to right would be marked as 3.1; 3.2; 3.3; 3.4 and so on depending on the number of letters published.

Editorial page follows the same criterion for paper identity. In the editorial page ‘middles’, quotes or historical columns from the past issues which are not regular columns, have been left out. The variables included for editorial pages are: themes, gender, geographic location. Location is further divided into: International, National, Regional. Urban, Rural, and Other category have been omitted for the editorial pages. As normally photographs are not published in the editorial page, photo prominence is excluded. The editorial page is more or less standardised in its format and layout, and content, so headline size and column weightage are not
studied. Gender is included in the theme category depending on whether the editorial or article is about Male, Female or Other.

Thus, in the editorial pages we have studied the following variables:

1. Paper Identity - items in the op-ed page follow PPCNN for paper identity
2. News Item Identity
3. Content Quality which examines whether the content is Responsible or not
4. Geographic Location to know whether the item refers to 1 International, 2 National, 3 Regional
5. Gender of the item to classify an item as Male (M), Female (F), Other (O)
6. Subject or Theme of the item decided on the basis of under which thematic category does the item occur and the frequency of the occurrence.

5.10 Computer Programme Used for the Study

A computer programme based on FoxPro was developed specifically for this study. All data entry and coding have been done using this software. Computerisation of data was considered necessary due to the large quantity of news and feature items we had to code and investigate as part of the present research.

The sample size of the study is 312 issues distributed between The Assam Tribune and The Shillong Times. Both these newspapers have 8 columns, each
column having several news items. Hence attempting to record the variable attributes of every news item under study with regard to prominence index, reporting index and news classification manually without the aid of a computer would indeed be a herculean task. Further, trying to analyse the data to gather meaningful insights by traditional manual counting techniques would render the task too difficult and time consuming. This is where computerisation becomes immensely helpful.

The recording of data is done covering a period of three years and involves a large amount of work. Consistency in the classification of data is an important factor to be kept in mind even when the work is spread out over a long period of data. A computer programme with built in checks and balances can take care of maintaining consistency in the task of data collection. It can therefore avoid subjectivity that is likely to creep in especially in the absence of a well-defined and rigorous classification built into the process of data collection. The computerisation of data collection has enabled the researcher to do away with large number of registers, notebooks and score sheets with elaborate tables that would have been required in the absence of a computer programme.

The computer programme had built-in controls to ensure that accidental mistakes would not creep in. The unique identity of each news item ensured that the same news item would not be analysed more than once. This may seem trivial; but it is a real possibility when the data collection is spread over several months. The menu system ensured that while an attribute of a news item would be recorded as a code number, the user did not have to memorise any code, but merely select a given description from a list of choices. It was also possible to add to the list of choices, as
for example in the case of subject themes, as it is not possible to make an exhaustive
list of themes at the start of the data collection process.

The programme was designed to enable the research to address the problem of
analysing all data during or after the coding. It took care especially of the following:

- Counting
- Correlating data
- Filtering – subsets
- Multiple filtering conditions

The greatest advantage of computerisation was obviously in the task of
analysing the data. It would be extremely difficult, tedious and error-prone to have to
scan a total of 24,450 news items every time some statistical data about their
attributes is required. The task becomes even more forbidding when one has to
correlate two or more attributes of the 24,450 news items. The computer
programme, instead is able to render the results with precision and accuracy in a
fraction of a second.

The computer programme also allows the researcher to do calculations on a
subset of the data collection, irrespective of the criteria used to filter out the subset of
items from the larger data set. Hence, the programme ensured speed, accuracy and
possibility of retrieval, if it became necessary for reassessment.

The computer programme used for data-entry was developed using FoxPro
9.0. This Windows-based package lent itself to creating user-friendly screens for easy
data entry. Adequate error-trapping procedures were incorporated to prevent
inadvertent mistakes in data-entry. This made it possible to avail the services of an
assistant to help in data entry while the researcher concentrated on the actual content and its classification under the various aspects like prominence index, reporting index and classification.

The attributes of the news items were recorded in tables having structures as detailed below.
Table 5.10 Structure for Table: CONTENT.DBF

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Type</th>
<th>Width</th>
<th>Dec</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PAPERID</td>
<td>Numeric</td>
<td>3</td>
<td></td>
<td>A unique identity for a paper.</td>
</tr>
<tr>
<td>2 NEWSID</td>
<td>Character</td>
<td>5</td>
<td></td>
<td>A unique identity for each news item</td>
</tr>
<tr>
<td>3 COLWT</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>The column weightage</td>
</tr>
<tr>
<td>4 HEADLINE</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>The headline weightage</td>
</tr>
<tr>
<td>5 NEWSLOC</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>The location weightage</td>
</tr>
<tr>
<td>6 PHOTO</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>Weightage for photo</td>
</tr>
<tr>
<td>7 SPACEALLOT</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>Space allotted</td>
</tr>
<tr>
<td>8 CONTENTPN</td>
<td>Character</td>
<td>1</td>
<td></td>
<td>Positive, Negative, Other (P/N/O)</td>
</tr>
<tr>
<td>9 CONTENTQ</td>
<td>Numeric</td>
<td>6</td>
<td>2</td>
<td>Content Quality</td>
</tr>
<tr>
<td>10 LOCHEOG</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>Location Geographic (1,2,3)</td>
</tr>
<tr>
<td>11 LOCUR</td>
<td>Character</td>
<td>1</td>
<td></td>
<td>Rural, Urban, Other</td>
</tr>
<tr>
<td>12 GENDER</td>
<td>Character</td>
<td>1</td>
<td></td>
<td>Male, Female, Other (M/F/O)</td>
</tr>
<tr>
<td>13 NEWSTYPE</td>
<td>Numeric</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.11 Structure for table: EDITPAGE.DBF

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Type</th>
<th>Width</th>
<th>Dec</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PAPERID</td>
<td>Numeric</td>
<td>3</td>
<td></td>
<td>A unique identity for a paper</td>
</tr>
<tr>
<td>2 ITEM</td>
<td>Numeric</td>
<td>4</td>
<td>2</td>
<td>A unique identity for each news item</td>
</tr>
<tr>
<td>3 CONTENTQ</td>
<td>Numeric</td>
<td>5</td>
<td>2</td>
<td>Content Quality</td>
</tr>
<tr>
<td>4 LOCHEOG</td>
<td>Numeric</td>
<td>1</td>
<td></td>
<td>Location Geographic (1,2,3)</td>
</tr>
<tr>
<td>5 GENDER</td>
<td>Character</td>
<td>1</td>
<td></td>
<td>Male, Female, Other (M/F/O)</td>
</tr>
<tr>
<td>6 NEWSTYPE</td>
<td>Numeric</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The figures in the next pages give two sample shots for data-entry input.

Figure 5.1 Sample Shot -1
Figure 5.2 Sample Shot-2
The analysis of the data and all the statistics obtained from it was done using FoxPro directly and the outputs of the queries were recorded in an Excel worksheet. A separate programme was not developed for statistical analysis. This was found to be convenient since any type of query could be handled and results obtained by anyone with sufficient knowledge of FoxPro.

Data gathered in the present research is an invaluable mine of raw information from which one is able to draw a lot of information pertaining to the research. As this rich resource of data and its interpretation have not been fully exhausted, there still remains the possibility of obtaining more information depending on the query and type of co-relation required. Further research and interpretation can yield other meaningful and validated conclusions.