Chapter 2
Literature Review: Content Analysis

Literature survey constitutes an integral part of any research in social sciences. In the present study we have examined the literature available on the subject of content analysis and social responsibility of the press. The literature reviewed here includes theoretical studies on content analysis and empirical research on projects working with qualitative content analysis and literature on computer programmes in support of qualitative content analysis and frequency word count. Pollock’s study of newspaper content using community structure approach as well as the impact study of newspaper content is of particular significance to the present study.

2.1 Content Analysis as a Research Method

The term Content Analysis as a research methodology is nearly a century old but it may rightly be admitted that its history dates back to human history itself, to the beginning of human use of symbols and language. Today content analysis is no longer defined by its traditional application of understanding meaning of messages. It has, over the years, developed into a methodology in its own right that it enables the researcher to plan, to communicate, and to critically evaluate a research design.
independently of its results. Krippendorff (1980) presents a brief history of content analysis and distinguishes it from other methods and exemplifies its domain of practical application.

As a research method content analysis uses a set of categorisation procedures for making valid and replicable inferences from data (text, voice or images) to their context. Inferences may be about the messages within the texts, the writer(s), the audience, and even the culture and time of which these are a part. Texts can be defined broadly as books, book chapters, essays, interviews, discussions, newspaper headlines and articles, photographs, historical documents, speeches, conversations, advertising, theatre, informal conversation, or any occurrence of communicative language. Since it can be applied to examine any piece of writing or occurrence of recorded communication, content analysis today is used in a variety of fields, ranging from marketing and media studies, to literature and rhetoric, ethnographic and cultural studies, gender and age issues, sociology and political science, psychology and cognitive science. Additionally, content analysis reflects a close relationship with sociology and psycholinguistics, and is playing a key role in the development of artificial intelligence.

Communication research is traditionally divided into two basic genres: Media-centric research which focuses on the structure of the communications industry and on media content, with minimal attention to impact on the intended audience. Effects-centric research focuses on audience effects and takes media organisation and content as given. Content analysis is traditionally practiced as a media-centric methodology.
However vast may be the study, content analysis requires that a particular collection of texts is initially selected as a sample. The identification of relevant features is always subject to individual interpretation in accordance with the aims of a particular study. Even if the purpose of a study is clearly defined, no two researchers would be likely to generate the same initial categories in a collection of texts. As Winick et al (1973) note, in a content analysis based on advertisements containing a wide assortment of tangible and intangible elements, content item definitions cannot be refined to a point where all subjectivity is eliminated from the coding process. It is not necessary to subscribe to the ideology of absolute objectivity in order to employ content analysis.

Content analysis, as noted earlier, is extensively used to study characteristics and impact of media on society and individuals. It is used to study a variety of content- newspapers, books, vision statements of institutions and organisations, policy documents of governments and non-governmental organisations (NGOs), police cases, criminal records, suicide notes and so on. Sinha (1980) has listed several other functions of content analysis such as identifying the intentions of the communicators, determining psychological states of persons and groups, detecting propaganda and subversive tendencies in the publications of organisations and groups, securing political and military intelligence, legal purposes, study of cultural patterns, social and individual values, mores, interests, attitudes and so on. Hart’s (1933) pioneering study of the contents of selected popular magazines in USA from 1900-1930 revealed changing social attitudes and interests.
Content analysis pertains not only to the study of the manifest content of the material. Becker and Lissmann (1973) have differentiated levels of content: themes and main ideas of the text as primary content and context information as latent content. The analysis of formal aspects of the material belongs to its aims as well. Content analysis embeds the text into a model of communication within which it defines the aims of analysis. This is expressed by Krippendorff (1969) who defines content analysis as "the use of replicable and valid method for making specific inferences from text to other states or properties of its source." Qualitative content analysis defines itself within this framework as an approach of empirical, methodologically controlled analysis of texts within the context of communication, following content analytical rules and step-by-step models, without rash quantification.

We can distinguish different phases in the historical background of content analysis (Merten, 1983; Krippendorff, 1980; Mayring, 1994a). Approaches to analyse and compare texts in hermeneutic contexts vary. These include Bible interpretations based on textual analysis, early analysis of newspaper content, graphological procedures, the dream analysis by Sigmund Freud. The basis of quantitative content analysis has been laid by Paul F. Lazarsfeld and Harold D. Lasswell in the United States during the twenties and thirties of the twentieth century. The first textbook about content analysis as a method was published by Berelson in 1952.

In the sixties of the twentieth century there was an interdisciplinary broadening and differentiation. Content analysis found its way into linguistics,
psychology (Rust, 1983) and other social sciences such as sociology, history and arts. The procedures have been refined (fitting into different models of communication); analysis of non-verbal aspects, contingency analysis, computer applications (Pool, 1959; Gerbner et al, 1969). Since the middle of twentieth century objections were raised against a superficial analysis without respecting latent contents and contexts, working with simplifying and distorting quantification (Kracauer, 1952). In the following years qualitative approaches to content analysis was further developed and perfected as a methodology for social science researches (Ritsert, 1972; Mostyn, 1985; Wittkowski, 1994; Altheide, 1996).

2.2 Definitions of Content Analysis

Krippendorff (1980) defined content analysis as a “research technique for making replicable and valid inferences from data to their context.” Berelson (1952) defined it as “a research technique for the objective, systematic and quantitative description of the manifest content of communication.” Berelson listed some seventeen uses of content analysis most of which pertain to study of communication. Stone & Dunphy (1966) saw journalism and mass communication as the historical origins of the technique of content analysis. Holsti (1969) placed data in the context of communication between sender and receiver.

The following list adapted from Berelson (1952) offers more possibilities for the uses of content analysis. Content Analysis can be used to:

- Reveal international differences in communication content
- Detect the existence of propaganda
• Identify the intentions, focus or communication trends of an individual, group or institution

• Describe attitudinal and behavioural responses to communications

• Determine psychological or emotional state of persons or groups.

The early definitions of content analysis were quite narrow and focused more on quantitative characteristics. Making inferences from the communication content is the primary function of content analysis. It does not study behaviour itself, but it infers data of communication. Content analysis has proved that 'text' is an excellent vehicle for studying long-term changes of attitude, concerns and styles (Sinha, 1980). Study of text provides insights into the context. Lasswell’s classical definition of communication as “who says what, to whom in which channel” (Lasswell, 1942) points to content. But in content analysis the study centres more on what is said than what is understood by the receiver. The significance of the content or what is said becomes one of the chief concerns of the present study. As what is said, in most instances cannot be fully abstracted from the way it is said, we also study the quality of what is said. In the present study, content analysis is both quantitative as well as qualitative. It involves measuring as well as judging and evaluating. The researcher here is concerned with the content as well as method, since any attempt to study communication in its entirety would require both these components.
2.3 Content Analysis as Quantitative and Qualitative Research

Content analysis is a standard methodology used in social sciences like communication and can be applied in quantitative and qualitative research. However, scholars on the subject like Holsti (1969), Krippendorf (1980) consider content analysis as a quantitative research and excludes the possibility of qualitative research. Veal (2006) and others consider it as a research methodology suited to quantitative and qualitative study of the data. Quantitatively it starts with word counts, space measurements (column centimeters in the case of newspapers), time counts (for radio and television). Qualitatively it may involve any kind of analysis where communication content is categorised and classified.

Qualitative research is the umbrella term given to a group of research methods applied to the collection and analysis of qualitative data or information. Rather than merely looking at the numerical data it stresses that the fundamental basis of qualitative research is that it is used to attempt an interpretation of reality at a certain point of time and within certain contexts. As the world is not a fixed phenomenon, but is interpreted by individuals in many different ways, it opens an opportunity for research into these different perspectives of the world.

Quantitative research deals traditionally with anonymous statistics, where people are numbers on a spreadsheet. The current research goes beyond what Kelly (1980) described as purely quantitative research as it deals with detailed analysis of qualitative data which are capable of yielding reliable results. This research takes an inductive approach in that it ultimately serves the purpose of identifying a range of
issues and recurring patterns in the data. The researcher then offers a detailed
description of the findings from the data analysis which leads to conclusions, as well
as creating a framework for issues to be discussed in further study (Merriam, 2000;
Peterson, 1994).

Content analysis can be quantitative or qualitative (Veal, 2006), the former
being where frequencies of a word or image are counted and placed within categories
to which a statistical analysis will be applied. The present research has adopted a
combination of interpretive quantitative and qualitative approach. It attempts to
investigate and interpret, within the context of the research subject, how the press
depicts socially relevant issues.

Qualitative content analysis wants to preserve the advantages of quantitative
content analysis for a more qualitative text interpretation. There are obvious
advantages. But qualitative content analysis should be determined on the basis of
what part of the communication inferences shall be made, to aspects of the
communicator (his/her experiences, opinions, feelings), to the situation of text
production, to the socio-cultural background, to the language and literary style, to the
text itself or to the effect of the message. The material is to be analysed step by step,
following rules of procedure, devising the material into content analytical units.

The aspects of text interpretation, following the research questions, are putted
into categories, which are to be carefully founded and revised within the process of
analysis (feedback loops), criteria of reliability and validity: the procedure has the
pretension to be inter-subjectively comprehensible, to compare the results with other
studies in the sense of triangulation and to carry out checks for reliability. For
estimating the inter-coder reliability, qualitative content analysis is used instead of quantitative content analysis.

Although the equation of content and meaning should be rejected, there can be no dispute that texts do have content and that the generation of meanings is not unconstrained by such content. As long as one avoids reductive assumptions about how such content will be interpreted by others, content analysis can provide useful data: for example, data about statistically significant differences between features in one set of texts compared with another; differences that may not be obvious from simple observation. Furthermore, the use of such a technique is not incompatible with associated qualitative studies of how individuals interpret the same texts. Since content analysis is a technique for a certain kind of textual analysis, the investigation of meaning (for instance, the sense that children make of advertisements) requires other techniques such as ethnographic interviews.

2.4 Procedures of Qualitative Content Analysis

The above listed components of quantitative content analysis will be preserved to be the fundament for a qualitative oriented procedure of text interpretation. Among the number of procedures of qualitative content analysis two approaches are central: inductive category development and deductive category application (Mayring, 2000). We shall examine what each of these categories mean.
i) Inductive Category Development

Classical quantitative content analysis has few answers to the question from where the categories come and how the system of categories is developed. Though very little is written about it, within the framework of qualitative approaches it would be of central interest, to develop the aspects of interpretation, the categories, as near as possible to the material, to formulate them in terms of the material. For that scope qualitative content analysis has developed procedures of inductive category development, which are oriented to the reductive processes formulated within the psychology of text processing (Ballstaedt et al, 1981; van Duk, 1980).

Step model of inductive category development (Mayring, 2000) is a procedure to formulate a criterion of definition, derived from theoretical background and research question, which determines the aspects of the textual material taken into account. Following this criterion the material is worked through and categories are tentative and deduced step by step. Within a feedback loop those categories are revised, eventually reduced to main categories and checked in respect to their reliability. If the research question suggests quantitative aspects (e.g. frequencies of coded categories) it can be analysed.

ii) Deductive Category Application

Deductive category application works with prior formulated, theoretically derived aspects of analysis, bringing them in consonance with the text. The qualitative step of analysis consists in a methodologically controlled assignment of the category to a passage of text. Even if several procedures of text analysis are processing that step, it is poorly described. Here we are referring to the step model
within qualitative content analysis (Mayring, 2000). The main idea here is to give explicit definitions, examples and coding rules for each deductive category, determining exactly under what circumstances a text passage can be coded with a category. Those category definitions are putted together within a coding agenda.

2.5 Practical Applications of Content Analysis

Content analysis has often been used as a tool for determining authorship of books and manuscripts when authorship is not known, or are in dispute, or when the work is very ancient. One technique for determining authorship is to compile a list of suspected authors, examine their prior writings, and correlate the frequency of nouns or function words to help build a case for the probability of each person's authorship of the data. Mosteller and Wallace (1964) used Bayesian technique based on word frequency to show that Madison was indeed the author of the Federalist papers. Foster (1996) used a more holistic approach in order to determine the identity of the anonymous author of the 1992 book Primary Colors. Content analysis has helped to establish the authorship of several books of the Bible. For instance, the Letter to the Hebrews in the New Testament of the Bible, formerly considered to be the work of Saint Paul, was found to be the work of some other writer than Paul. By analysing the language of the text and literary style, allusions and other internal evidences deduced from the work, it is also possible to establish the period in which such a text was written.

Content analysis is also useful for examining trends and patterns in documents. For example, Stemler and Bebell (1998) conducted a content analysis of
school mission statements to make some inferences about what schools hold as their primary reasons for existence. One of the major research questions was whether the criteria being used to measure programme effectiveness (e.g., academic test scores) were aligned with the overall programme objectives or reason for existence. It is also possible to compare stated goals and objectives of institutions with specific tasks adopted to achieve them. The method not only helps to understand language competence and literary standard, but also to examine motives, personality traits, character, bias, worldviews - all of which find expression in the way language is used.

Further, content analysis provides an empirical basis for monitoring shifts in public opinion. Data collected from the mission statements project in the late 1990s can be objectively compared to data collected at some point in the future to determine if policy changes related to standards-based reform have manifested themselves in school mission statements. Shift in policy and perspective can be studied if election manifestoes of political parties in various elections over a period of time are studied through this method. Content analysis is used to examine suicide notes, anonymous letters, crime records, nature of authority etc. Investigators of crime have often been able to establish whether a death is a suicide or a murder by analysing the contents of the so-called ‘suicide notes’ and by verifying its authorship.

2.5.1 Conducting a Content Analysis

According to Krippendorff (1980), one of the leading scholars on content analysis, six questions must be addressed in every content analysis. They are:

1. Which data are analysed?

2. How are they defined?
3. What is the population from which they are drawn?

4. What is the context relative to which the data are analysed?

5. What are the boundaries of the analysis?

6. What is the target of the inferences?

At least three problems can occur when documents are being assembled for content analysis. First, when a substantial number of documents from the population are missing, the content analysis must be abandoned. Second, inappropriate records (e.g., ones that do not match the definition of the document required for analysis) should be discarded, but a record should be kept of the reasons. Finally, content analysis should be abandoned if some documents might match the requirements for analysis but are uncodable because they contain missing passages or ambiguous content (GAO, 1996).

2.5.2 Analysing the Data

Perhaps the most common notion in qualitative research is that a content analysis simply means doing a word-frequency count. The assumption made is that the words that are mentioned most often are the words that reflect the greatest concerns. While this may be true in some cases, there are several counterpoints to consider when using simple word frequency counts to make inferences about important issues. Synonyms may be used for stylistic reasons throughout a document and thus may lead the researcher to underestimate the importance of a concept (Weber, 1990). Besides each word may not represent a category equally well since nuances and context can vary with each instance. Further, words may have multiple meanings. For instance the word ‘state’ could mean a political body, a situation, or a
verb meaning 'to speak'. The word 'spring' may mean one of the seasons, a source of water, a mechanical devise, or the act of springing in the air and so on. The nuances of words become clearer in specific contexts and situations in which they are used and not merely in the frequency with which they occur. As there are no well-developed weighing procedures, using word count requires the researcher to be aware of the many limitations it entails. Furthermore, as Weber (1990) reminds, "not all issues are equally difficult to raise".

A good rule of thumb to follow in the analysis is to use word frequency counts to identify words of potential interest, and then to use a Key Word In Context (KWIC) search to test for the consistency of usage of words. Most qualitative research software, such as NUD*IST, Hyper RESEARCH, allow the researcher to pull up the sentence in which a particular word was used so that the researcher can see the word in some context. This procedure will help to strengthen the validity of the inferences that are being made from the data. Certain software packages (e.g., the revised General Inquirer) are able to incorporate artificial intelligence systems that can differentiate between the same words used with two different meanings based on context (Rosenberg et al, 1990). There are several software packages available that can help facilitate content analyses.

However, content analysis extends far beyond simple word counts. What makes the technique particularly rich and meaningful is its reliance on coding and categorising of the data. The basics of categorising can be summed up in these quotes: "A category is a group of words with similar meaning or connotations" (Weber, 1990). "Categories must be mutually exclusive and exhaustive" (GAO,
Mutually exclusive categories exist when no unit falls between two data points, and each unit is represented by only one data point. The requirement of exhaustive categories is met when the data language represents all recording units without exception.

2.5.3 Coding Data

There are two approaches to coding data that operate with slightly different rules, namely, Emergent vs. A priori coding: With emergent coding, categories are established following some preliminary examination of the data. The steps to follow are outlined in Haney et al (1998) and may be summarised here. First, two people independently review the material and come up with a set of features that form a checklist. Second, the researchers compare notes and reconcile any differences that show up on their initial checklists. Third, the researchers use a consolidated checklist to independently apply coding. Fourth, the researchers check the reliability of the coding (a 95% agreement is suggested; .8 for Cohen's kappa). If the level of reliability is not acceptable, then the researchers repeat the previous steps. Once the reliability has been established, the coding is applied on a large-scale basis. The final stage is a periodic quality control check.

When dealing with a priori coding, the categories are established prior to the analysis based on some theory. Professional colleagues agree on the categories, and the coding is applied to the data. Revisions are made as and when necessary, and the categories are tightened up to the point that maximises mutual exclusivity and exhaustiveness (Weber, 1990).
2.5.4 Coding Units

There are several ways of defining coding units. The first way is to define them physically in terms of their natural or intuitive borders. For instance, newspaper articles, letters, or poems all have natural boundaries. The second way is to define the recording units syntactically, that is, to use the separations created by the author, such as words, sentences, or paragraphs. A third way to define them is to use referential units. Referential units refer to the way a unit is represented. For example a text might refer to Mahatma Gandhi as "the Father of the Nation," "Gandhiji" or "Bapu". Referential units are useful when we are interested in making inferences about attitudes, values, or preferences. A fourth method of defining coding units is by using prepositional units. Prepositional units are perhaps the most complex method of defining coding units because they work by breaking down the text in order to examine underlying assumptions. For example, in a sentence that would read, "Investors took another hit as the stock market continued its descent," we would break it down to: The stock market has been performing poorly recently/Investors have been losing money (Krippendorff, 1980).

Typically, three kinds of units are employed in content analysis: sampling units, context units, and recording units. Sampling units will vary depending on how the researcher makes meaning; they could be words, sentences, or paragraphs. In the present study the sampling unit is a particular 'story' or piece, which may be a news item, article, editorial, column, or letter to the editor, in general called 'stories' or 'items'. Context units neither need to be independent or separately describable. They may overlap and contain many recording units. Context units do, however, set
physical limits on what kind of data one is trying to record. In the present study, context units are stories or items. Recording units are rarely defined in terms of physical boundaries.

2.5.5 Reliability

Weber (1990) notes: "To make valid inferences from the text, it is important that the classification procedure be reliable in the sense of being consistent. According to him, "reliability problems usually grow out of the ambiguity of word meanings, category definitions, or other coding rules". Yet, it is important to recognise that the people who have developed the coding scheme have often been working so closely on the project that they have established shared and hidden meanings of the coding. The obvious result is that the reliability coefficient they report is artificially inflated (Krippendorff, 1980). In order to avoid this, one of the most critical steps in content analysis involves developing a set of explicit recording instructions. These instructions then allow outside coders to be trained until reliability requirements are met. Reliability may be discussed in terms of stability, or intra-rater reliability. Reliability can be ensured if the same coder is able to get the same results again and again.

2.6 Computer Programmes in Support of Qualitative Content Analysis

Computer applications have greatly enhanced the quality and speed of different social research studies, including content analysis of a variety of texts, images and data. During the last few years several computer programmes have been
developed within the framework of qualitative analysis to support (not to replace) steps of text interpretation (Huber, 1992; Weitzman and Miles, 1995; Mayring, 1996; Fielding and Lee, 1998). The computer plays here a triple role: it works as assistant, supporting and making easier the steps of text analysis on screen, doing such functions as working through the material, underlining, writing marginal notes, defining category definitions and coding rules, recording comments on the material etc. It offers helpful tools in handling the text and carrying out such processes as searching, jumping to different passages, collecting and editing passages etc.

It works as documentation centre, recording all steps of analysis of all interpreters, making the analysis comprehensible and replicable (e.g. to trace back the material in case of non-reliabilities between two coders). It offers links to quantitative analysis (often already implemented within the programme), e.g. to compare frequencies of categories, without the danger of errors in data transfer by hand to another computer programme.

Hansen (1995) provides an overview of using information technology to analyse newspaper content and to access and analyse electronic texts of newspaper stories. He discusses the use of word processing, concordance, and text retrieval software and reviews some of the basic approaches to coding newspaper content. Kelle (1996) in his paper on “Computer-Aided Qualitative Data Analysis: An Overview” describes how the latest generation of software for qualitative data analysis (e.g., NUDIST, HyperResearch) enables researchers to create nested and linked codes. He also reviews the hypothesis testing features of these programmes. He suggests that these features are perhaps more appropriate for quantitative rather
than qualitative analysis, although he notes that hypothesis testing can be used profitably by qualitative researchers provided they exercise caution in interpreting their results.

2.7 Content Analysis in Journalism

While journalism is a formal discipline in academics, content analysis emerged as an instrument for analysing journalism nearly a century ago. Willey’s (1926) study of newspapers, *The Country Newspapers* in 1926, is one of the earliest sociological analyses of newspaper content.

Some of the earliest applications of content analysis were related to quantitative analysis of printed materials. Loebl (1903) published in German an elaborate classification scheme for analysing the ‘inner structure of content’ according to the social functions newspapers perform. Hotchkiss and Franken (1920) published a study of the newspaper reading habits of college students. They sent a questionnaire to 1761 men and women college students in New York. A similar study was done by W D Scott in 1916 in Chicago. He sent the questionnaire to four thousand prominent people in business and industry. These studies revealed that on an average people spent only a quarter of an hour a day to read newspapers.

In 1910, Max Weber (1911) proposed a large-scale content analysis of the press but for a variety of reasons it did not get off the ground. Since then with the growth of journalism schools and proliferation of newspapers and periodicals, the quest for quantitative assessment and ethical standards have grown. Quantitative newspaper analysis arose as a response to this quest. Speed (1893) in one of the
earliest studies of content analysis, raised the rhetorical question, “Do newspapers Now Give the News?” The author showed how religious, scientific, and literary matters had dropped out of leading New York newspapers between 1881 and 1893 in favour of gossip, sports and scandals. A similar study attempted to reveal the overwhelming space devoted to ‘demoralising’, ‘unwholesome’ and ‘trivial’ matters as opposed to ‘worthwhile’ news items (Mathews, 1910). By simply measuring the column inches a newspaper devotes to particular subject matter, journalists attempted to reveal “the truth about newspapers” (Street, 1909). Content can help find a way of showing the profit motive as the cause of “cheap yellow journalism” (Wilcomx, 1900). Fenton (1910) became convinced that the study of content established “the influence of newspaper presentations on the growth of crime and other antisocial activity”. White (1924) concluded that a “quarter-century survey of the press shows demand for facts”.

Quantitative newspaper analysis contains many valuable ideas. In 1912 Tenney called for a large-scale and continuous survey of press content to establish a system of bookkeeping of the “social weather” comparable in accuracy to the statistics of the U.S. Weather Bureau. Sociologist Willey’s (1926) study of The Country Newspaper traced the emergence of Connecticut country weeklies, reported their circulation figures, changes in subject matter, and the social role they acquired in competition with large city dailies. Quantitative newspaper analysis was used in response to questions like how the Negroes were presented in the Philadelphia press (Simpson, 1934). Allport and Fender (1940) studied newspaper content from an entirely new perspective. Their “Five Tentative Laws of the Psychology of
Newspapers' attempted to account for the changes that information undergoes as it travels through an institution and finally appears on the printed page.

Content analysis is used to identify, evaluate and audit standards, to establish what kind or how good an object is. It has been used for examining trends and patterns. Speed (1893) compared several New York dailies published in 1881 with those published twelve years later and observed changes in subject matter categories. Lasswell (1942) proposed a study and presented preliminary findings on trends in the frequencies with which references to various countries occur in different national presses. Another significant factor in content analysis is examining the differences in communication. The two newspapers, namely, *The Assam Tribune* and *The Shillong Times*, the contents of which are being analysed in the present research, obviously have several differences such as ownership, geographical location, the manner of covering news, content, policy, type of readers, priority etc. In these two newspapers we can discern a close relation between differences and trends.

Hoover Institution's study of 19,553 editorials that appeared in American, British, French, German, and Russian prestige papers during 1890-1949 led Pool (1959) to observe-predict that proletarian doctrines replace liberal traditions, that an increasing threat of war is correlated with growth in militarism and nationalism, and that hostilities toward other nations are related to insecurity.

The Royal Commission on the Press in England appointed by the Queen in 1947 made use of mainly content analysis of subject and form from the newspapers, besides other things, to ascertain the extent of suppression, distortion and inaccuracy over a period of twenty years (Sinha, 1980). Bettleheim and Janowitz (1950) studied
racial, ethnic prejudice by means of unstructured interviews and by analysing the content of their responses. Content analysis has been widely used to measure the suitability of printed textbooks, to study stylistic features, language and literature. In Northeast India, content analysis has been used by a few researchers to study academic curriculum and textbooks. However, content analysis as a methodology has not been used to carry out systematic and scientific study of print or other media in Northeast India.

The evaluation of the performance of the press has been a major preoccupation since the emergence of quantitative newspaper analysis. The concern with changes from quality to quantity of news reporting (Speed, 1893) or with increase of trivial, demoralising, and unwholesome subject matter at the expense of worthwhile information (Mathews, 1910) assumes evaluative standards, albeit, implicitly. It is difficult to agree on a sufficiently acceptable scale that places newspapers between good and bad. Many of the evaluative studies have limited themselves to measuring biases. Even this becomes difficult, as there is no consensus on what constitutes journalistic bias, impartiality etc. Merrill (1962) used a battery of evaluative criteria on journalistic presentations (attribution bias, adjective bias, adverbial bias, contextual bias, photographic bias, and outright opinion) but the catalogue is far from complete.

Many content analyses use measurable entities as indices of not so directly measurable phenomena. To cite a few examples: speech disturbance ratio to measure a patient’s anxiety during a psychiatric interview (Mahl, 1959), the frequency of certain words indicating motive to be achieved (McClelland, 1958), typographical
space as an index of importance to an issue (Budd, 1964), television violence index (Gerbner et al, 1979), indices of citizen dissatisfaction computed from letters of complaint (Krendel, 1970).

Content analysis has more specifically been used to study propaganda. The first significant study of propaganda was by Kris and Leites (1947) who analysed propaganda in the two world wars. The most glaring example of the use of propaganda is by German Nazi leader and politician, Joseph Goebbels (1897-1945) who was Hitler’s minister of Propaganda from 1933, during the Nazi regime. Goebbels’s task was supposed to have been equivalent to today’s minister for communications and information. He has given us the word ‘Goebbelesim’ to mean propaganda and misinformation. It includes manipulation of information, unethical practices, lies, deception, misinformation by the political powers, often using the press, for purposes such as political advantages, commercial interests, or other acts of irresponsibility.

Since then we have numerous instances of misuse of information, as during the Vietnam war spearheaded by the then US administration, and the more recent instance of manipulating information during the Gulf war and the invasion of Iraq and the overthrow of Saddam Hussain. We have today the term ‘embedded journalism’ to mean manipulative practices in journalism in which the journalist becomes an agent of the government or the powers that control society. In India too, manipulation of information content is used by political parties during the elections. Political parties continue to manipulate information in various ways to suit their needs and to achieve certain ends.
The press by engaging in any manipulative method of using information belies the trust placed on it by the public. One recent and glaring example is that of *The Shillong Times*, dated April 1, 2007, which misled its readers by publishing four fake stories on its front page and in the sports page on April 1, 2007. This was to coincide with the April Fools’ Day. The next day the newspaper carried a disclaimer. This kind of gimmicks was continued on April 1, 2008 as well. Many readers resent such moves by newspapers, which already suffer from lack of credibility or have a fast dwindling readership and circulation. Many readers were angry for duping and misleading them by indulging in distasteful April Fools’ Day pranks. News and information is manipulated and planted in the pages of newspapers also by the various militant organisations in Northeast India. They use the media to suit their propagandist goals.

As every system and ideology tend to use information to suit the preservation and promotion of its objectives, content analysis can provide guidance and direction to readers, and reduce misuse and manipulation of information. No serious and systematic study of newspaper content has been undertaken in India, and hence, such studies are necessary to ensure a healthy relationship between the press and the public in a country that is so vast and diverse. Media content has, time and again, inflamed passions and led to communal flare-ups and riots. The disruption of the Valentine’s Day celebrations in parts of the country, the vandalisation of the paintings of India’s leading artist M.F. Hussain, forbidding the shooting of certain films, the publication of the novel, *Satanic Verses*, of Salman Rushdie are a few instances that illustrate the point.
2.8 Frequency and Word Count

Frequency of a symbol, word, idea and subject indicate importance, attention, emphasis given to that issue. The count of favourable or unfavourable attribute of a symbol or idea helps to interpret or measure bias. Lasswell (1942), Leites and Pool (1942) emphasise the importance of studying word as an important unit of content analysis. Though word count and frequency of words are considered important elements in content analysis, in the present study the researcher has not adopted word count or frequency of occurrence of particular words as a unit for quantitative analysis or as a unit of measurement. He has instead used the qualitative analysis to suggest context, to enable easy interpretation, to deduce meaning and value judgment. While word count and word frequency are suitable to measure attitudes or biases related to specific themes like aggression, violence, peace, war propaganda it is difficult to apply it effectively when the entire newspaper issue is analysed for several quantitative and qualitative characteristics. Excessive and exclusive dependence on word count could in certain instances mislead and provide erroneous inferences. For instance, Pool’s (1951) observation that symbols of democracy appear less frequently when a representative form of government is accepted than where it is in dispute suggests the deficiency in mere word count and frequencies without considering their correlation with other issues. It further points to the fact that words do not exhaust communication content. Silence, signs, symbols, body language and other non-verbal elements too communicate meaning.

Word count and frequency devoid of the context can yield misleading results. For instance, in a country ruled by a dictatorial government, the word ‘democracy’
might occur in the newspapers more frequently than where democracy may be functioning effectively, as in the former, there could be protests against curtailment of democracy, frequent demands for introduction or restoration of democracy. The difficulty becomes more compounded when one has to use a computer programme to count word frequency. The computer can mechanically carry out word count without reference to the context in which it is used or without being able to interpret the nuances of the words. Such applications of content analysis can fail to provide contextual meanings, connections and linkages between words and ideas. Word count and word frequencies are indicative not only of quantity but must also take note of qualitative correlation and measurements in order to obtain a comprehensive idea of the content in context and its meaning.

2.9 Examples of Projects Working with Qualitative Content Analysis

There are several examples of empirical studies using qualitative content analysis procedures: Vicini (1993) conducted 14 open-ended in-depth interviews with educational advisors about concrete case-studies from their advisory service with the aim of reconstructing the theory of mind of advice. He used summarising qualitative content analysis leading to eight main categories. The results were that advice praxis had become therapy-oriented, that there are totally different concepts of advice.

Gerwin (1993) made a diary study with 21 middle school teachers about their daily hassles and uplifts and analysed the transcripts with summarising qualitative content analysis. She could demonstrate that being a teacher means severe stress,
from everyday problems with the copy machine to treating students with behaviour disorders.

Beck and Vowe (1995) analyzed 25 media products (newspapers, journals, radio transmissions) concerning new multimedia approaches. They used a combination of inductive and deductive qualitative content analysis. They found patterns of argumentation like: euphoria about multimedia; economic optimism; political critic; apocalyptic predictions.

Dolde and Goetz (1995) have conducted five open-ended interviews with adult students in an on-the-job computer education studio. Working with inductive and deductive qualitative content analysis they analysed their learning activities and learning strategies. The main advantage of the learning concept in the course seemed to be flexibility in time, as main disadvantage appeared heterogeneity of course members.

Bauer et al (1998) analyzed the biographies of 21 Alzheimer disease patients to find out common patterns and to compare them with 11 vascular dement patients of the same age. The biographical interviews had been worked through with qualitative content analysis and led to typical biographical patterns (e.g. overprotecting social network) of the Alzheimer patients.

Mayring et al (2000) studied unemployment of teachers in the eastern parts of Germany. The material had been collected through open-ended interviews and open-ended biographical questionnaire of 50 unemployed teachers, asking for their psychosocial stresses and coping behaviour. The results were compared with a former study on teacher unemployment in West Germany by the same team. Inductive and
deductive computer-assisted content analysis pointed out that the doubled crisis situation of the persons (unemployment and German unification) caused specific stress and new chances for adaptation.

With the qualitative content analysis Mayring et al (2000) wanted to describe procedures of systematic text analysis, which try to preserve the strengths of content analysis in communication science (theory reference, step models, model of communication, category lead, criteria of validity and reliability) to develop qualitative procedures (inductive category development, summarising, context analysis, deductive category application) which are methodologically controlled. These procedures allow a connection to quantitative steps of analysis if it seems meaningful for the analyst. The procedures of qualitative content analysis seem less appropriate, if the research question is highly open-ended, explorative, variable and working with categories would be a restriction, or if a more holistic, not step-by-step ongoing of analysis is planned. On the other hand qualitative content analysis can be combined with other qualitative procedures. The research question and the characteristics of the material should have the priority in making decisions about adapted methods. Mayring opines that it is better to discuss questions about methods in respect to specific content areas (Mayring, 1994b; Schmitt & Mayring, 2000) and then to compare different methodological approaches, including quantitative approaches.

A host of scholarly studies which examine communication content are available. Evans et al (1995) in “Science Content and Social Context" reviewed content analytic studies of popular science news and recommend that researchers do
more to examine visual imagery and non-elite news content. They also recommend that content analysts incorporate perspectives from linguistics and rhetoric and more formally contextualise content analytic data in terms of social theory. Brosius et al (1995) in their article "Killer and Victim Issues: Issue Competition in the Agenda-Setting Process of German Television" challenge the "equal-displacement" model of agenda setting in which it is assumed that when media coverage of an issue increases, media attention to other issues decreases more or less evenly across issues. Instead, the authors develop a "restructuring" model in which one can identify "killer" issues that are relatively more likely to displace other issues and "victim" issues that are relatively more likely to be pushed off the media agenda by "killer" issues.

Gonzenbach (1996) carried out a Longitudinal Analysis of the Drug Issue in which he presented the interactions between media coverage of illegal drugs, presidential structuring of drug issues, government policy regarding illegal drugs, public concern regarding illegal drugs, and real-world measures of the severity of the illegal drug problem. He identified four stages of media coverage (pre-problem, discovery, plateau, and decline) and reports that public concern and media attention lead presidential and government attention to the issue of illegal drugs.

Ideodynamic model of the relationship between media coverage and public opinion developed by Hertog et al (1995) was applied to newspaper and newsmagazine coverage of HIV transmission. News coverage of whether or not HIV can be transmitted via toilets, sneezing, or insects was found to predict public belief (as measured through survey research) in the likelihood of HIV transmission via these routes.
Popovich (1995) reviewed quantitative magazine research published between 1983 and 1993, including analyses of magazine advertising, photo coverage, minority coverage, and other content features. Finding a relatively small number of studies of magazine audiences and effects, he concluded that "magazine researchers rely too heavily on content analysis techniques, remain fragmented in focus, and lack theoretical foundations from which to pose their research questions."

According to Craig (1995) there are limitations to the traditional agenda setting models. He suggests an alternative approach that relies on longitudinal analysis and considers the interplay of news coverage of an issue, scientists' attention to an issue, government attention to an issue; and public concern with an issue. This approach allows researchers to identify distinctive phases in the agenda setting process and to more rigorously identify interactions between variables.

Randall (1995) examined the development of quantitative measures of verbal intensity and reported on an experiment that showed that verbal intensity scores generated with an elaborated definition of intensity were more highly correlated with subjects' estimates of the intensity of discourse manifested in audio recordings of discourse than were scores generated with an earlier definition. He elaborated definition of verbal intensity, assessed the frequency of death references, qualifiers (or 'intensifiers'), obscure words, sex references, profanity, and metaphors.

Textual ambiguity is the subject of study of Roberts, Carl W and Popping (1996). They drew attention to three potential sources of textual ambiguity that may challenge researchers seeking to construct network representations of texts. Texts may contain idiomatic ambiguity (i.e., the presence of idiomatic words and phrases),
Illocutionary ambiguity (in which the illocutionary intention of a sentence is unclear),
and relevance ambiguity (in which some relationships between themes may be
irrelevant). They contended that researchers who conduct network analyses of texts
must do more to take into account these three sources of ambiguity.

Ben (1997) developed and tested a theory that international conflict can be
predicted by assessing the number and nature of a nation's newspaper editorials that
focus on other nations. The author reported that these measures successfully predict
international conflicts around the world over the past several decades. He suggested
that mass media are crucial in government attempts to cultivate public consensus that
conflict with other nations is justified.

The article of Kabanoff (1996) aimed at organisational behaviour researchers
provides an overview of computer-supported content analysis, arguing that
documents generated by organisations can provide valuable data about organisational
behaviour, including data about organisational goals and the cognitive processes of
top managers. He discussed the advantages and limitations of computer-supported
content analysis and reviews several organisational behaviour studies that have made
good use of computer-supported content analysis.

In many content analyses, only a sample of all coded units are coded by
multiple coders. Lacy, Stephen and Riffe (1996) contend that this method of
assessing intercoder reliability poses potential problems that have seldom been
discussed in the extant literature. The authors provide a formula designed to estimate
sample sizes required for valid reliability assessments. In addition, they discuss the
need for more careful reporting of reliability assessment procedures when these procedures involve sampling of coding units.

In a study of Russian political speeches delivered between 1964 and 1993 Anderson, Ronald D Jr (1996) reported that Russian politicians manifested a decrease in the ratio of nouns to verbs, a decrease in the use of the conjunction "and" relative to "but," a decrease in the use of the negative prefix ("un" and its variants) relative to the negative particle ("not" and its variants), an increase in the use of first-person singular voice, and an increase in the use of personal pronouns. The author claims that these changes reflect a decreasing conceptual distance between politicians and the populace. Such changes, the author suggests, are typical of shifts from authoritarianism to democracy.

Mergenthaler, Erhard (1996) argues for a ‘marker view’ of lexical content analysis in which words are seen as indicators of the presence of themes rather than as direct measures of themes. He describes his computer-assisted system for identifying cycles and key moments in transcripts of clinical psychological sessions. This system works by charting the frequency of words that have been shown to be correlated with emotional expression and abstraction.

Perloff (2000) in his study records how during the 1930s, after thousands of African Americans had been put to death by mobs - particularly in the South, but in other regions of the country as well - lynchings were no longer unusual or shocking events that deviated from the norm. Although there have been many studies of racial biases in the modern media and a host of scholarly investigations of the African American press during the late nineteenth century, Perloff (2000) says that there has
been virtually no research examining the ways in which the mainstream American press covered the lynching epidemic that swept the South during the late nineteenth and early twentieth centuries. “In view of the paucity of research, it is not surprising that journalism history textbooks devote virtually no space to press coverage of lynchings. Many papers in southern parts of U.S.A. provided vicious coverage of lynchings during the late nineteenth century.”

Drawing on historical works, secondary sources, and hundreds of newspaper accounts, Perloff (2000) summarises what we know about how newspapers discussed lynching of Blacks in America on their news and editorial pages during the nineteenth and early twentieth centuries. Far from suppressing news about lynchings, newspapers embraced them, providing abundant, even graphic, coverage of vigilante violence. As Clark (1964) observes in a book on the southern country editor, "many editors did not spare their readers' sensibilities. Whatever their motives, they (editors) wrote full, detailed accounts.” Most lynching stories contained grizzly details, identified the victim as a Negro, and even suggested that the accused was guilty of the crime and therefore deserving of punishment (Perloff, 2000). Some newspapers and magazines denounced the practice of lynching Black Americans. Articles on lynchings of the Blacks had a special vitriolic quality. Newspaper stories identified the race of the accused, assumed without question that the accused person was guilty, used a number of dehumanising terms to label the Black victim, E.g., ‘wretch’, ‘fiend’, and ‘desperado’, assumed the Black person's race predisposed him to commit violent crimes, particularly rape, and sometimes self-righteously defended lynching of Black individuals. Newspaper data on lynchings became very helpful for
researchers to obtain quantitative facts about press biases. Studying the press coverage of lynching can shed light on the way various kinds of biases affect newspaper reporting.

2.10 Pollock’s Study of Newspaper Content Using Community Structure Approach

John C. Pollock has used content analysis to study US newspapers on a variety of social and ethical issues. His research has demonstrated the usefulness of content analysis as an effective tool to measure contemporary issues ranging from newspaper coverage of Islam post 9/11, ethical and moral issues like cloning, homosexuality, Internet privacy etc. In all the studies reviewed below, Pollock has used community structure approach.

2.10.1 Nationwide Newspaper Coverage

i) Islam Post-September 11

In a nationwide study of nineteen US newspapers, Pollock (2005) tracked the coverage of Islam post- September 11. The study, undertaken a year after the attack on the twin towers of the World Trade Center in New York on September 11, 2001, explored links between city characteristics and coverage frames by combining ‘prominence’ and “direction” scores to yield a single ‘Media Vector’, measuring issue projection. Using Lexis Nexis, this study explored 325 articles, all of 500 words or more in nineteen major US newspapers printed between September 11, 2001 and September 12, 2002. All articles were coded and assigned two scores – ‘prominence’ score, a numerical rating from 3 to 16 points based on placement of the article in a
newspaper, headline size, length of the article, and the presence of photographs or graphics. A second score called ‘directional score’ was used to determine coding each article’s ‘framing’ as favourable, unfavourable, or neutral/balanced. Most (fourteen) cities revealed positive coverage. Pearson correlations and regression analysis yielded several significant ‘stakeholder’ characteristics negatively correlated with favourable coverage of Islam. Contrary to expectations that a greater presence of foreigners, in particular Arabic or Farsi speakers, would be linked to appreciative perspectives, the higher the percentage of foreign-born citizens or number of Arabic/Farsi speakers the less favourable the coverage of Islam. In our research, we have used Pollock’s concept of ‘prominence score’ with modifications. We have termed it ‘prominence index’.

Pollock’s study examined how a national cross section of major US newspapers reacted to critical choices about framing Islam after September 11 and how much the content of their coverage reflected the communities they serve. He concluded that while newspapers are active in setting public agendas, they are also influenced by the society that surrounds them.

ii) Homosexuals in the Boy Scouts of America

Using a ‘community structure approach’, Pollock et al (1977, 1978, 1994-2002), carried out a nationwide study and explored systematic links between community demographic characteristics and newspaper reporting on homosexuals in the Boy Scouts of America (BSA), especially the Supreme Court decision permitting gay exclusion. All articles over 150 words on the topic published between January 1, 1998, and March 1, 2001, were selected from a national cross-section sample of
twenty-one newspapers, yielding 322 articles. A single score, the Media Vector, was calculated to combine article ‘prominence’ as well as reporting direction (favourable, balanced/neutral, or unfavourable).

**iii) Privacy on the Internet**

In this study, primary research was conducted to investigate how newspapers from different cities in the United States reported on Internet Privacy. The study compared hypotheses on different city characteristics and newspaper coverage of Internet Privacy using the ‘community structure approach’, tested in earlier versions in Minnesota by Tichenor, Donohue and Olien (1973, 1980) and elaborated in nationwide studies by Pollock and others (1977, 1978,1994-2000), suggesting that certain demographic structures of a community are systematically linked to newspaper reporting on critical issues.

A national cross-section sample of twenty-one newspapers was selected from the DIALOG newspaper database, and a sample of 25 articles above 350 words in each newspaper on Internet Privacy were drawn from January 27, 1996 to November 7, 2000. The resulting total of 495 articles was analysed using content analysis. A single score, the Janis-Fadner Coefficient of Imbalance, was calculated to combine attention given to each article as well as reporting direction (favourable, balanced/neutral, or unfavourable). Pearson correlations were used to link different city characteristics to coverage of Internet privacy regulation.

**iv) FDA Regulation of Tobacco**

The report on FDA Regulation of Tobacco discussed the way in which newspapers in major cities across the United States have covered the issue. The
community structure approach developed by Tichenor, Donohue, and Olien (1973, 1980), and elaborated by Pollock and others (1977, 1978, 1995, 1996, 1997, 1998), suggests that the structures of communities, regarding certain demographics in particular, are linked systematically to the coverage of critical issues in its newspapers. The approach was used to create and test a set of hypotheses studying the correlation between city characteristics and newspaper reporting on FDA regulation of tobacco.

A sample of articles with a minimum of 400 words was drawn from newspapers in fifteen major cities using the DIALOG database. They were all written between the years 1993 and 2000. The final total of 272 articles was then analysed, focusing on content and direction. The content was based on article placement, headline size, word count, and direction. The direction (favourable, unfavourable, or balanced/neutral) was then assigned to each article and used to calculate the Janis-Fadner Coefficient of Imbalance for each newspaper. The results showed that newspaper coverage varied widely throughout the nation, with coefficients of imbalance ranging from .246 to -303.

v) 1998 Tobacco Master Settlement Agreement

The 206-billion dollar tobacco industry agreed to pay 46 states in November 1998, in an unprecedented act. Previously, although sued several times, the tobacco industry was almost always victorious. When former tobacco company employees became willing to testify, these outcomes changed. The tobacco industry was shown to have intentionally stalled scientific and health awareness regarding smoking throughout the 50s. States sued to recoup health care money spent on ill smokers.
Ultimately the Master Settlement Agreement (MSA) was developed as a legal remedy, signed by 46 states, joining Florida, Minnesota, Mississippi and Texas, which came to similar settlements a few months earlier.

This study explores the connection between city demographics and nationwide news coverage of this event, specifically using a 'community structure approach' initiated in Minnesota by Tichenor, Donohue and Olien (1973, 1980) and elaborated in nationwide studies by Pollock and colleagues (1977,1978, 1994-2001). This approach suggests that particular demographic structures of a community are systematically linked to newspaper reporting perspectives on critical issues.

vi) Physician-Assisted Suicide

Physician-assisted suicide has recently been a topic of vigorous debate, yet few studies seek to explore this highly charged topic as a communication issue. Unlike other studies exploring the impact of media on society, this investigation examines the impact of society on media, specifically linking city characteristics to systematic content analysis of newspaper coverage of physician-assisted suicide. Specifically, this study maps the way newspapers from a national cross-section of cities across the United States differ in their coverage of physician-assisted suicide. The community structure approach suggests that certain demographic structures of a community are systematically linked to newspaper reporting on critical issues. This approach was used to test a set of hypotheses exploring the correlation between city characteristics and nationwide newspaper coverage of physician-assisted suicide.

The study, part of a continuing series exploring the relation of city characteristics to newspaper coverage of "critical events" such as, Magic Johnson's
announcement, Dr. Kevorkian's activities or tobacco's Master Settlement Agreement, confirms the strong association nationwide between community structure and media alignment with political and social change.

vii) Embryonic Stem Cell Research

This study uses a "community structure approach" to explore the connection between city characteristics and nationwide newspaper coverage of embryonic stem cell research. A sample of 350 articles, chosen from twenty-one newspapers across the nation, was coded for 'prominence' and article 'direction' (favourable, unfavourable, or neutral). The results were combined to calculate a single-score 'Media Vector' for each newspaper. Pearson and regression analysis revealed that three variables accounted for 85% of the variance: "health care access" (number of physicians per 100,000 residents); "stakeholders" (% Catholic and % Republican); and "media access" (% cable-subscribers). Healthcare and media access are linked to favourable coverage of stem cell research.

viii) Detainee Rights

The terrorist attack on the World Trade Center and the subsequent war in Afghanistan have created some unique problems for the United States government. One of the issues to be dealt with is what to do with the former Taliban fighters who have been captured in Afghanistan. The most dangerous have been transported to the detainee camp in Guantanamo Bay, Cuba, where they are being held under heavy guard indefinitely. A controversy arose over how to treat the detainees. The issue not only created a partisan split, it polarized parts of the Bush administration and put the U.S. at odds with the rest of the world.
This study uses a “community structure approach” to explore the connection between city characteristics and nationwide newspaper coverage of extending Prisoner of War status to the Taliban and Al Qaeda detainees imprisoned at Guantanamo, Cuba. A sample of all the longest articles over 150 words in a nationwide cross-section of twenty-two newspapers in the sample period January 1 to March 15, 2002, yielded 302 articles. These were coded for “prominence” and article “direction” (favourable, unfavourable, or neutral) and combined to calculate Pollock’s “Media Vector” for each newspaper. The results confirmed Pollock’s “stakeholder” hypothesis. Contrary to expectation, Pearson correlations revealed significant relationships between voting Republican and favorable detainee rights coverage. By contrast, higher percentages of Catholics, Democrats, and Hispanics were found significantly related to negative coverage of detainee rights. Regression analysis revealed that five variables accounted for 95% of the variance: “stakeholders” (% Republican); “media access” (% cable-subscribers); “buffer” (% Income over US$ 100,000); “Position in Lifecycle” (families with children aged between 8-18); and “violated buffer” (% Crime).

ix) Music Censorship

A nationwide study of music censorship compared hypotheses linking different city characteristics and nationwide newspaper coverage, using the “community structure” approach to research, as developed in nationwide studies by Pollock and others (1977, 1978, 1994-2002).
x) Same-Sex Adoption

Same-sex adoption is an emerging controversy. In some states same-sex adoption is supported, but in conservative Florida, homosexual couples are forbidden from adopting children. Yet the American Academy of Pediatrics welcomes same-sex adoption. Supporters of gay adoption argue that since there is no conclusive evidence on whether children of gay families are negatively impacted, and also because many children who are products of homosexual parents are happy and well-adjusted, there is no reason to ban same-sex adoption. Opponents of same-sex adoption argue that both male and female counterparts are needed to adequately raise a family and homosexual parenting raises the issue of sexual abuse. Differing opinions surrounding this issue make same-sex adoption an important topic to explore.

Specifically, this study uses a community structure approach, initiated by Tichenor, Donahue, and Olien in Minnesota (1973, 1980) and revised and tested by Pollock and other researchers in nationwide studies (1977, 1978, 1974-2003) to explore the ways certain demographic structures of a community are linked to newspaper reporting on same-sex adoption. A sample of newspaper articles of 150 words or more printed on the topic was collected from Lexis-Nexis from October 31, 2000 to December 1, 2002. The resulting 216 articles were analysed using content and statistical analyses. Content analysis combined a “prominence” score -- comprised of the article’s placement, length, headline size, presence of photographs or graphics -- and overall “direction” -- favourable, unfavourable, or neutral/balanced -- for each article to calculate Pollock’s “Media Vector” for each newspaper, yielding a measure of issue “projection.”
2.10.2 City Characteristics and Newspaper Coverage

i) NAFTA

Since many citizens identified the North American Free Trade Agreement (NAFTA) with employment/unemployment issues, Pollock et al (Pollock, 1995) used this critical event as an ideal opportunity to study the association between citizen/influential concerns and newspaper coverage. Using DIALOG, a national newspaper database, he analysed locally generated articles published from September 1, 1993 to December 31, 1993, twenty major newspapers representing a geographic cross-section of cities in the United States. A content analysis technique was used to evaluate both the "amount of attention" an article received and its "direction" to yield a single score for each newspaper. These scores were compared with a variety of city characteristics to test hypotheses associating several aspects of community structure with reporting variations.

Employing correlation and multiple regression analyses, one key factor was found clearly associated with coverage favouring or opposing NAFTA. Contrary to positions taken by organised labour, the higher the proportion of the labour force employed in manufacturing in a city, the more positive that city's reporting on NAFTA is likely to be. Mapping coverage of the NAFTA debates demonstrates that archival data comparing newspaper databases and city characteristics can reveal significant variations in reporting on political and economic policies.

ii) Human Cloning

In February 1997 Dolly, a cloned sheep, and her maker, Scottish scientist Ian Wilmut, hit the headlines. The story figured prominently in newspapers, on
television, across the Internet, and in conversations. Since the ability to clone humans became a reality, there was a media frenzy sparking many a debate among a variety of publics: scientists, lawyers, ethicists, religious leaders, government representatives and others. Pollock et al (2000) tracked news coverage from different regions of the United States sampled systematically in twenty-two newspapers throughout the nation during the period of January 1997 through December 1998. The time frame included both coverage on Dolly, the first sheep cloned, and Richard Seed, the person who announced he had the ability to begin cloning humans. Previous studies suggest that variations in community or city characteristics (using aggregate data and demographics) have a great deal to do with variations in reporting on critical issues).

The DIALOG Classroom Information Program, a national newspaper database, was used to collect 380 articles over one paragraph in length. Each of the articles was read and coded for two kinds of information: a "display" or "attention" score (ranging from 3 to 16 points based on article placement, headline size, article length, and presence of photographs, captioned or not) and a "directional" score (legitimizing/favourable, delegitimising/unfavourable, and balanced/neutral).

iii) Magic Johnson HIV Announcement

Although many articles have appeared in the past few years on the connection between mass media outputs and perceptions of HIV/AIDS, only recently has attention been drawn to the relation between reporting on HIV/AIDS and the structural characteristics of local communities. A recent "white paper" drafted by distinguished health communication scholars for the NIMH Office of AIDS asks that researchers begin to explore links between national or community-specific norms and
reporting patterns. This research (Pollock et al, 1994) is one of the first systematic efforts to test that linkage by examining a cross-sectional national sample of newspapers. Over 450 articles were studied using a content analysis technique that ultimately assigns a single "score" to each of twenty-one papers in distinct U.S. geographic regions. Comparing the rankings of those scores with city rankings on a variety of selected aggregate data indicators, several findings emerge.

Four clusters of hypotheses were tested, regarding size (city population and media saturation); belief system differences; athletic and fitness involvement; and status and achievement -- privileged lifestyles. Curiously, there is little relation between variations in coverage of Johnson and such city characteristics as: ethnic identity (percent Black or Hispanic); fan enthusiasm (percent television sports fans); or belief system variations (percent Catholic; percent engaged in devotional reading). Income matters relatively little. In sum, a community structure analysis of newspaper coverage suggests that the Magic Johnson HIV announcement reflects achievement, success and fitness issues more than ethnic or moral issues, or even sports fan involvement. The apparent Johnson message: hard work and success carry no immunity from a deadly virus that is voraciously democratic.

The above studies led by Pollock deal with a variety of contemporary themes. They show content analysis of newspapers can be used today with great degree of scientific accuracy to measure and examine how newspapers and other media deal with issues of relevance to the community at large. Such studies are easily replicable and are particularly useful in measuring newspaper’s objectivity, social responsibility, bias and a host of other attitudes that can affect the content.
2.11 Impact Study

The Readership Institute of the Media Management Center at Northwestern University, USA, in 2001 embarked on the largest content analysis of newspapers ever completed. This study, which measured newspaper content, is named the 'Impact Study'. It examined 100 US newspapers, analysing 47,000 stories in 700 issues across the United States (Lynch and Peer, 2002). Rather than asking consumers to describe what kind of news and information they want to know and how they should be covered, the Impact Study measured newspaper content, measured consumer reaction and then looked for significant correlations between the two. 37,000 readers and non-readers in the 100 Impact newspaper markets reported how important different types of information and news are to them. They were also asked how satisfied they were with their local newspaper's coverage of these topics. The study pertained to five standard circulation ranges. Newspapers were ranked in size from 10,000 to 600,000 average daily circulations. The Readership Institute developed its own tools to measure content and trained a team of researchers.

The researchers measured news content in those same local newspapers, documenting exactly to what consumers were reacting. The Readership Institute adopted a multi-dimensional approach to measure each story which helped to gauge what types of news are covered as well as how news is covered. Of the 74,000 stories in the selected papers, 47,000 stories were chosen for in-depth analysis. The Readership Institute designed a multi-dimensional approach to measure each story.
For each story the researchers measured several characteristics including the theme, origin (source) geographic focus, news style, visual complexity, front page diversity, overall counts and content organisation.

The study examined the percentage of coverage of various kinds of news and found that the smaller papers give more emphasis on local stories on the front page than the larger papers, which give greater emphasis to national and international stories on the front page. The study also examined how news is covered – photographs, graphics, difference on the coverage on Sundays, gender diversity in sourcing. A 450-question survey among 37,000 readers was also conducted. The study revealed that even the slightest increase in overall content satisfaction increases readership. In other words, the study revealed what content practices related to consumer satisfaction (transcripts of interviews, discourses, protocols of observations, video tapes, documents etc).

### 2.12 The Press in India

Though the press has played a very important role in the political and socio-cultural history of India, media research is quite new and scanty. As much of the studies available in India are based on media research done in the West, there is great scope and vital need for media research in the country. Today with the mushrooming of several media faculties and university departments on the one hand, and a growing and vibrant press and electronic media, media research assumes great significance. Content analysis, as we have seen, can be an import research method for serious study of media content in general of the press in particular.
One significant study of newspaper content regarding the way press covers issues pertaining to women was undertaken by Ammu Joseph and Kalpana Sharma (1994). They studied newspaper coverage of five major issues - dowry, rape, Shabano case, foeticide and Sati. The period covered was from 1979 to 1988. The data was both quantified in terms of the total number of items on each issue and their break-up across separate categories like news, editorials, articles etc. The qualitative analysis was done by looking at the placement of the item within the publication and analysing its content. The placement and type of coverage given to an issue signify the relative importance granted to it by the newspaper. Of the 498 items related to the five themes, Sati topped the list with 41.1 per cent, and Shabano case was the second with 29.1 per cent. Five leading English newspapers were chosen for the study, namely Hindustan Times, The Indian Express, The Times of India, The Statesman and The Hindu. The study was carried out on four leading vernacular language press-namely, Bengali, Gujarati, Hindi and Tamil. The study was also extended to the development of women’s programming on Indian television and the ideology of some of its ‘women oriented serials telecast in the mid-eighties.

Much of the literature available on the press in India pertains to the history and growth of the press. Ever since the Portuguese missionaries established the first printing press in Goa in 1556, the press in India has played an important role in the growth of literacy, education, development of languages, and the promotion of social consciousness. Many of the pioneers of newspapers and periodicals in the country had clear nationalistic goals. As we have seen in the previous chapter, the press served as an important instrument in the struggle for India’s independence. Besides
the Europeans, who included missionaries, several Indian leaders like Raja Rammohun Roy, Bal Gangatarak Tilak, Mahatma Gandhi, used the print media with the social and nationalistic goals in mind. (Sharma, 1994)

There is an abundance of literature on the pioneering years of the press in India and other south Asian countries. However, most of these are not related to critical studies of content, but rather deal with the history and development of the press. Lent's (1981-82) scholarly article ‘The Missionary Press of Asia 1550-1850” outlines the history of the press in Asia. Natarajan (2000) traces the history of the press in India. The press in India has not only survived the onslaught of those who tried to muzzle and suppress it but became one of the powerful forces the authorities had to reckon with in the struggle for independence. Natarajan chronicles this struggle and highlights the important milestones in the growth and development of the press in India.

There are several other scholarly and historical studies on the press in India. We may mention the most outstanding of these: Aurobindo Mazumdar (1993) deals specifically with the role of the press in the freedom struggle from 1937-42. Parthasarathy (1991) brings in more than four decades on his experience to tell the story of the press in India and the doyens of the Indian press. Sarkar (1984) presents the history of the press in India from the legal perspective. Bhargava (2005) has traced the evolution of the press in India as an industry and the way it has coped with the challenges and changes that marked various eras. Kesavan (1988) in a two-volume study presents a comprehensive history of printing and publishing in India.
The history of the press in India and the freedom struggle go hand in hand. Not only did the press powerfully assist the country’s aspiration and yearning for freedom, but it has also, in the post-independence era, stood firmly in upholding the democratic principles enshrined in the Constitution of India.

A free press has played the role of a ‘watchdog’, a handmaid to democracy, and a responsible ‘Fourth Estate.’ India has been able to maintain its democratic system of government uninterrupted for more than six decades while several other Asian countries, formerly under colonial powers, had to live with periods of authoritarian governments. In addition to the wisdom of the framers of India’s Constitution and the sagacity of the people of this nation, we might pay tribute to an independent and responsible press, which has contributed in no small measure to the social and political process that shaped and continues to shape the destiny of India.

But one can witness winds of change everywhere. In this first decade of the 21st century there is widespread apprehension about the survival of newspapers. We are witnessing an alarming decline of newspaper circulation and readership worldwide. In many countries such as India the picture is quite different. The print media in India is far from being dead, if the official report on the state of the printed publications is anything to go by. According to Press in India 2004-5, the annual report of the Registrar of Newspapers for India (RNI), the print media claimed a substantial share in the information space in the country registering 1,948 new newspapers and over a two crore increase in circulation in 2004-05. As on March 31, 2005, there were 60,413 registered newspapers on record as against 58,469 at the end of March 2004. With an annual increase of 2,36,31,621, the total circulation of...
newspapers in India in 2004-05 was 15,67,19,209 copies. Four newspapers also ceased publication during this period.

2.13 The Press in Northeast India

History of the press in northeast India dates back to the establishment of the first Assamese monthly *Orunodhoi* in 1846. The region’s press is hence 162 years old. (Plathottam, 1996) The press in Northeast India has played the role of a catalyst, not only merely to inform, educate or entertain, but also to promote ethnic and communal harmony, nationalism, public opinion and other democratic functions. The geographic features of the region, the presence of several languages, and the ethnic diversity have had a decisive impact of the press in the region.

Data on the number of newspapers in the region indicates an ever-growing, vibrant press in the region. Print media in the region has made significant progress during the last fifteen years. Today there are 1074 registered titles with Registrar of Newspaper for India (RNI) as compared to 700 at the beginning of 1992. Hence we can observe 53% growth during this period. Meghalaya is leading in the number of registered periodicals with a growth rate of 78%, followed closely by Assam with 67%. In Assam, out of 15 daily newspapers registered in Assamese language, 12 have been started during the last fifteen years (Bharali, 2007).

While statistics indicate rapid growth of the press in Northeast India, there has been no scientific study of the content and social responsibility of the press in the region. Hence it is our attempt to document, study and evaluate the contents of two of the leading newspapers of the region, *The Assam Tribune* and *The Shillong Times* and present evidence in support of conclusions on the social responsibility of the press in
northeast India. The thesis is an attempt to provide empirical evidence in support of the concept of social responsibility of the press in northeast India.

2.14 Social Responsibility

Social responsibility implies that the press assumes voluntarily the task of providing adequate coverage to all sections of the community, be they economically or ethnically significant or powerful. In this section we intend to review literature on laws, legislation and guidelines enacted in various countries pertaining to the social responsibility of the press. These include provisions enacted for a free and socially responsible press in the Constitutions of various countries, journalism reviews started by well-known journalism institutes which critically examine the role and function of the press, Press Councils, Commissions, Editors’ forums and independent organisations concerned with fair and responsible practices in journalism.

Since the publication of *A Free and Responsible Press* by the Hutchins’ Commission on Freedom of the Press in USA in 1947 there has been a renewed interest in the moral dimensions of journalism (Schmuhl, 1989). There exists a close and intimate link between journalistic ethics and social responsibility of the press. Without defining and specifying ethical standards it would be difficult to lay down the ground rules for a socially responsible press. Even though journalism schools place a great deal of importance to ethics in their curriculum, rarely do we find journalists referring back to them in their journalistic practices.

Grover (1990) traces the press laws and the emergence of institutions like the Press Council. One of the first countries in Europe to legislate on the press was
Sweden, which established the Press Fair Practices Commission in 1916. Washington Journalism Review established in 1961, and Columbia Journalism Review, which followed, paved the way for strengthening awareness about the social functions of the press. An Editor's Note column in New York Times started in 1983 by A.M. Rosenthal, Executive editor, examined significant lapses of fairness, balance or perspective. Other attempts along similar lines were 'Inside Story' on PBS, 'Viewpoint' on ABC, Mellet Fund for a Free and Responsible Press established in the late 1960s, the National News Council, established in 1973, which was wound up in 1984, and a host of other initiatives.

In 1985, the UNESCO asked the international and regional organisations of professional journalists to prepare a review (Nordenstreng, 1989) of developments concerning the status, rights and responsibilities of journalists around the world. This work coordinated by Kaarle Nordenstreng, President of the International Organisation of Journalists, and ably assisted by a Turkish expert Hifzi Topuz, was undertaken on behalf of several leading press organizations in various continents like the International Organization of Journalists (IOJ), International Federation of Journalists (IFJ), Federation of Latin American Journalists (FELAP), Federation of Latin American Press Workers (FELATRAP), Union of African Journalists (UAJ), Federation of Arab Journalists (FAJ) and Confederation of ASEAN Journalists (CAJ). Though the editors of the volume admit that the work is far from being a comprehensive review, it is the first time at the international level such a well-documented publication was brought out. The work is rich in data, case studies, facts and figures, documents and viewpoints. The UNESCO commissioned study under the
leadership of Sean MacBride (1980) and published under the title *Many Voices, One World*, more popularly called the MacBride Report, calls for a New World Information, Communication Order (NWICO).

In India, the social responsibility of the press can be traced to the role the press played in the freedom struggle, the writings of Mahatma Gandhi and the other leaders involved in the freedom struggle, the Constituent Assembly and its deliberations on the press and its responsibilities, the post-independent legislations of the government and other bodies like the Press Council of India, the Press Commissions, the Advertising Standards Council of India etc. Bhabani Bhattacharya (1969) in his book *Gandhi the Writer*, presents the spirit that moved Gandhi to write extensively. He sheds much light on Gandhi's journalistic activities. K.E. Eapen (1995), one of the pioneers of journalism and communication education in India, calls for serious and systematic training in journalism. S.C. Bhatt's (1997) book, *Indian Press Since 1955*, is a well-researched work on the growth of the press. Besides these we have a large corpus of literature on the press by pioneers of the Indian press and journalism, the publications of Press Institute of India, books and reviews published by other media institutes in the country. *Vidura*, the journal of the Press Institute of India in its July-September 2006 issue dealt with the theme: 'Media and Social Responsibility.' The Right to Information Act enacted by Indian government in 2005 provides greater possibility for the press to seek information and offers easier access to sources of information of a public nature. The Act can help enhance a more effective exercise of social responsibility.
In this section we have reviewed the literature on content analysis and issues related to quantitative and qualitative factors that are related to the process of content analysis. Several computer programmes available today have made content analysis easier and more reliable. We have also reviewed some of the major projects working with content analysis. We have reviewed John C. Pollock’s studies of newspaper content using community structure approach to understand social issues.

We also reviewed the Impact Study by Readership Institute, USA, which examined 100 US newspapers, analysing 47,000 stories on 700 issues across the United States. The Impact Study is perhaps the largest content analysis of newspapers ever completed. We have included in the review of literature materials on social responsibility as well as literature on the press in India and Northeast India.

Research in newspaper content is gaining a great deal of importance today as newspapers are undergoing rapid changes and are facing serious challenges from the new media. The availability of newer tools to aid social research and the establishment of many media institutes of research, one can hope that content analysis would be able to advance further as a tool for media research.