

CHAPTER 3

CHAPTER 3 - METHODOLOGY

This chapter presents a description of the methods, techniques and empirical measures employed to collect and analyze the data. It has been organized and presented as follows.

3.0 Context

3.1 Description of population

3.2 Sampling method

3.3 Sample profile

3.4 Design of the study

3.5 Variables

3.6 Construction of tools

3.7 Reliability and validity of tools

3.8 Procedure followed for data collection

3.9 Period of the study

3.10 Statistical Analysis

3.0 Context

The main objective of the study was to ascertain the growth of online newspapers and their implications on print media. As online journalism in India is

still in its infancy, an exploratory research was taken up to study the likely impact of online newspapers on print media.

To examine the impact in its totality, the researcher further categorized the study into:

1. Impact on newspaper industry
2. Impact on journalists
3. Impact on content or journalism in general
4. Impact on the users

3.1 Description of population

All the leading newspapers which were published from Bangalore or which had editions coming from Bangalore were taken. They included, English dailies, *Deccan Herald*, *The Times of India*, *The Hindu*, *The New Indian Express*, *Asian Age*, and Kannada dailies, *Samyukta Karnataka*, *Vijaya Karnataka*, *Udayavani*, *Prajavani*, *Kannada Prabha* and evening *Sanjevani*. All these constituted the population for the editors' survey.

Population for the journalist survey included, all the journalist working in the online desks of print newspapers and those who were working for purely online

newspapers and portal sites, in three languages viz. Kannada, English and Marathi, as the survey was conducted in Bangalore and Bombay, capital cities of Karnataka and Maharashtra respectively.

Internet users in Bangalore who browsed in cyber cafes, offices and home PCs formed the population of the users' survey. Workplace covered academic institutions and offices where the users were students and non students. Cyber café was a common place which provided ready access to the public user for a fee based connect time. The population of the study also included people who owned computers and had internet connection and browsed at home.

3.2 Sampling method

In case of newspaper editors, 11 leading newspapers having online editions at the time of the study were contacted. Editors of all these newspapers were approached for the purpose.

In case of journalists, a disproportionate random sampling method was used to collect the required data from the population of the journalists mentioned above.

To collect the data of the users which was limited to Bangalore city, simple random sampling method was used.

3.3 Sample profile

For editors' survey the sample included the editors of *Deccan Herald*, regional English daily, and national newspapers like, *The Times of India*, *The Hindu*, *The New Indian Express*, and *Asian Age*. The Editors of Kannada dailies, *Samyukta Karnataka* which had started but temporarily stopped its online editions, *Vijaya Karnataka* which was about to launch its online edition at the time of the study, *Prajavani*, *Udayavani*, and *Kannada Prabha* which already had *online editions* and the only Kannada evening paper on the Net, *Sanjevani* were also approached for the details.

Of these, the editors of *Prajavani* and *The Times of India* could not be contacted for various reasons. Hence the responses of the editors of only nine newspapers have been taken for the final analysis.

The sample in case of journalists' survey included online journalists who were working for purely online newspapers, or on online desks of print

newspapers or were content managers in portal news sites, in Bangalore and Bombay.

The description of the respondents in case of user survey, is given in the form of table below. Their demographic details like sex, age, educational qualification, occupation and monthly income are also presented in the following table.

Table 3.1 : Distribution of respondents according to age educational qualification, occupation and monthly income versus sex

Sl No.			Male	%	Female	%	Total
1	Age	15 – 20yrs	16	16	12	24	28
		21 – 30yrs	39	39	29	58	68
		27 – 32yrs	26	26	5	10	31
		33yrs & Above	19	19	4	8	23
		Total	100		50		150
2	Education	S.S.L.C	4	4	4	8	8
		P.U.C	10	10	5	10	15
		Graduate	61	61	24	48	85
		Postgraduate	25	25	17	34	42
		Total	100		50		150
3	Occupation	Student	17	17	21	42	38
		Govt. Employee	2	2	0	0	2
		Pvt. Sector Employee	62	62	26	52	88
		Self Employed	10	10	1	2	11
		Unemployed	9	9	1	2	10
		Housewife	0	0	1	2	1
		Total	100		50		150
4	Monthly Income (Rs.)	3001 – 8000	56	56	18	36	74
		8001 - 13,000	23	23	21	42	44
		13,001 - 18,000	14	14	7	14	21
		18,001 and above	7	7	4	8	11
		Total	100		50		150

3.4 Design of the study

In the present investigation, *Ex-post-facto* research design was used. This was considered appropriate because the phenomenon had already occurred. The study was categorized into three parts to analyze the impact of online newspapers on print media, journalists and users.

To study the impact on newspaper industry a semi-structured interview was conducted with the editors of various newspapers cited above. The interview method was chosen because, elaborate data concerning respondents' opinions, values, motivations, recollection of experiences and feelings could be obtained and also allowed for lengthy observation of respondents' non-verbal responses.

To study the impact on journalist and to meet the objective working lives of online journalists, a survey was done using a questionnaire which was administered to online journalists in Bangalore and Bombay, who were, working for online editions of newspapers or known portal sites or purely online newspapers.

Along with this, an extensive survey of the Internet users in Bangalore was also done using another questionnaire specially designed for the purpose, to study the impact on ultimate consumers of information.

Questions regarding the impact on journalism as a whole were included in all the three types of surveys and were analyzed.

3.5 Variables

To study the over all impact of online journalism on print media a three-tier study was conducted with two different questionnaires administered to journalists and users or ultimate consumers of online newspapers. Based on the review of literature, discussions with experts and the objectives set for the study, the variables selected were categorized into two groups, namely; dependent and independent variables for all the three surveys.

In case of *editors' survey* only two *independent variables* were considered

1. **Size of the newspaper:** The size of the newspaper meant, the circulation of the newspapers, and their area of popularity, based on which the newspapers were divided as big or small.

2. **Language of the newspapers:** The sample newspapers were divided into language newspapers and English newspapers.

The *dependent variables* considered in this study were :

1. Average number of hits the websites received.
2. Organizational structure.
3. Technological know how.
4. Advertisement revenue.
5. Opinion on the future of online editions

For journalists' survey the following *independent variables* were considered.

1. **Sex:** It referred to the gender of the respondents, categorized as *male* and *female*.

2. **Age:** Age referred to the *age in years* mentioned by the respondents at the time of investigation. They were categorized into

- i) 21 - 25
- ii) 26 - 30
- iii) 31 - 35
- iv) 36 and above

3. Educational qualification: Educational qualification referred to the extent of formal education or schooling by the individual respondent. This question was left open because the new generation news rooms have persons from various faculties coming in. They were divided as those

- i) with Journalism education and
- ii) Without journalism education

4. Experience: Experience referred to the number of months or years worked in the field of journalism. This was further divided into two categories.

- i) Total journalism experience and
- ii) Online journalism experience

The *dependent variables* considered in this study were,

1. Technological know how of the journalists
2. Duties of the online journalists
3. Usage of the interactive options
4. Comfort level in the online environment
5. Opinion on the future of online journalism in India

In case of the *user's survey*, five *independent variables* were used.

1. **Sex** as defined in the journalists' survey.
2. **Age** as defined in the journalists' survey. But the respondents were categorized into
 - i) 15-20
 - ii) 21-26
 - iii) 27-32
 - iv) 33 and above
3. **Educational Qualification** as defined in the journalists' survey. The respondents were categorized into
 - i) S.S.L.C
 - ii) P.U.C
 - iii) Graduate
 - iv) Post-graduate

4. **Occupation:** It referred to formal profession or employment of the respondents. On the basis of their response, the respondents were categorized into

- i) Student
- ii) Government employee
- iii) Private sector employee
- iv) Self - employed
- v) Unemployed
- vi) Housewife

5 Monthly income: Monthly income referred to the income of the family of the respondent from all sources in *Indian rupees*, per month. This was categorized into

- i) 3001 - 8000
- ii) 8001 - 13000
- iii) 13001- 18000
- iv) 18001 and above

The *dependent variables* considered for the users' survey were

1. Readership of print newspapers
2. Ownership of PCs and Internet connection
3. Quantum of use of Internet
4. Purpose of Internet use
5. Newspaper readership on Net
6. Opinion on the advantages and disadvantages of Net as a medium for journalism

3.6 Construction of tools

In order to examine the overall impact of online newspapers on print media, specific objectives were chalked out. Keeping in mind the objectives, two different questionnaires were constructed with both closed ended and open ended questions to obtain the data. To understand the working lives of online journalists a questionnaire containing 25 questions with 19 closed ended and six open ended questions was constructed. For user survey, a questionnaire containing 20 closed ended questions was devised. For editors' survey a semi-structured interview was

conducted. All the three instruments were scrutinized by a panel of six judges, two of them being communication and journalism teachers, a former Chief Editor of a leading portal site, a journalist working for the online edition of a print newspaper, an online journalist working for a reputed portal and finally a senior statistician. The questionnaires were finalized based on the judgment of the expert panel. Information so accumulated had a very statutory purport in executing research task.

In order to test the reliability of the questionnaire a pre-test for users' survey was conducted with 30 questionnaires administered among Internet users drawn from the population consisting of men, women, home and workplace users, student and non student users, and coming from different socio-economic background. During this phase, an attempt was made to assess the objectives already determined; in order to ascertain whether the research action envisaged; is in fact realizable with the target sample in terms of its receptivity, motivation and interest, and was found to be compatible. Reliability test was then conducted on the pre-test data collected.

3.7 Reliability and validity of tools

The following data shows the results of *Split Half Reliability Test* of the *users' survey questionnaire*.

Table 3.2 : Table showing the results of split half reliability test of the users' survey Questionnaire

Name of the Tests	Reliability Coefficients
Correlation between forms	0.9240
Equal length Spearman-Brown	0.9605
Gutman Split half	0.7785
Unequal length Spearman-Brown	0.9456
Alpha for 1st part	0.4244
Alpha for 2nd part	0.4562
Alpha	0.6079

Validity = Square root of Guttman Split half i.e. $\sqrt{0.7785} = 0.8823 = 88.23\%$

Table 3.3. : Table showing the results of the *ANOVA* for the reliability analysis of the *users' survey questionnaire*.

Source of variation	Sums of Squares	df	Mean Square	F	Probability
Between Subjects	8247.80	29	284.41	17.5675	0.0000 Significant.
Within Subjects	254440.90	1470	173.09		
Between Items	95987.30	49	1958.92		
Residual	158453.60	1421	111.51		
Total	262688.70	1499			

The reliability of the scale was found to be 0.7785. This value was found to be sufficiently significant and the correlation between the two halves of the scale 0.9240 also acclaimed the reliability of the scale. The alpha for the scale was 0.6079, with alpha for first part being 0.4244 and for second part being 0.4562.

The reliability analysis was further subjected to ANOVA test which showed a significant difference between people and between measures. Significant F value of 17.5675, and probability 0.00 was indicative of the reliability of the measuring scale.

Square root of the Guttman Split Half 0.7785, i.e. 0.8823 showed that, the validity of the scale stood at an impressive 88.23 per cent.

At the same time as many as 50 e-mails were sent to online journalists working in various newspapers and portals in Bangalore and Bombay. Among those who responded 30 properly filled up questionnaires were taken, and reliability and validity test was conducted.

Table 3.4 : Table showing the results of the *Split Half Reliability test* of the *journalists' survey questionnaire*.

Name of the Tests	Reliability Coefficients
Correlation between forms	0.7689
Equal length Spearman-Brown	0.8693
Gutman Split half	0.8054
Unequal length Spearman-Brown	0.8572
Alpha for 1st part	0.4571
Alpha for 2nd part	0.4432
Alpha	0.5805

Validity = Square root of Guttman Split half i.e. $\sqrt{0.8054} = 0.8974 = 89.74\%$

Table 3.5 : Table showing the results of the *ANOVA* for the reliability analysis of the *journalists' survey Questionnaire*.

Source of variation	Sums of Squares	df	Mean Square	F	Probability
Between Subjects	9210.60	29	317.6069	19.6104 Significant	0.0000
Within Subjects	258227.00	810	318.7988		
Between Items	104174.00	27	3858.296		
Residual	154053.00	783	196.7471		
Total	267437.60	839			

The results of the Guttman Split Half reliability test of 0.8054 showed that the instrument used for the purpose was highly reliable. The correlation between the two halves of the scale 0.7689 also indicated the reliability of the scale. The alpha for the scale was 0.5805, 0.4571 for first half and 0.4432 for the second part.

ANOVA test which was used to confirm the reliability of the scale measured a significant difference between people and between measures. F value of 19.6104 with probability 0.00 further proved the reliability of the measuring instrument.

Square root of the Guttman Split Half 0.8054, i.e. 0.8974 showed that the validity of the scale was 89.74 per cent.

3.8 Procedure followed for data collection

To collect the data from the editors of various newspapers, appointment was fixed in advance, and a semi-structured, detailed interview was conducted with the editors. As the discussion progressed minute details of their problems in maintaining an online edition etc. were acquired and recorded carefully by the researcher in the *Field Visit Diary*. Nine editors were contacted for the purpose.

In case of journalists, the researcher personally approached as many online journalists as possible in the newspaper offices and content departments of portal sites in Bangalore and Bombay and administered questionnaires to them. Questionnaire was sent as an attachment to e-mails to journalists who preferred to answer online. Though the target was to reach 50 online journalists, the reluctance on the side of journalists made it possible to have opinions of only 47 journalists.

Questionnaires for users were administered in various cyber cafes of Bangalore which were selected from different parts of Bangalore to cover all socio-economic strata. Questionnaires were distributed in different offices and among college students. 150 properly filled questionnaires were taken for the final analysis.

3.9 Period of the study

The pre-test was undertaken in the month of June, 2002 and the actual data collection from the respondents was done from July to November 2002. There were obvious reasons for taking such a long duration, as the main difficulty was to get in touch with the high profile editors and online journalists. As there was fierce competition among the different newspapers websites and portal sites, the researcher was met with suspicion and there was initial reluctance to answer questions.

3.10 Statistical Analysis

The collected data was edited, coded and analyzed using Statistical Package for Social Sciences (SPSS). Percentages, simple tables and cross tabulations were used for the presentation of data. In consultation with a senior statistician, statistical techniques like, chi square test, t-test, ANOVA etc. were used to identify the relationships between different variables and finally the impact of online newspapers on print medium.

1. Percentages: The percentages were worked out to represent the proportion of respondents to the total sample considered in the study. This was calculated as

frequency of a particular group multiplied by 100 and divided by the total number of respondents in that respective category and in some cases the total number of respondents.

2. Chi-square test: At its basic, it is a non-parametric test often used to compare observed frequencies of a phenomenon with frequencies that might be expected or hypothesized. The researcher must interpret these data in a way that permits a statement of, whether the change in frequency is actually significant. This test is used for calculating the statistical significance between two proportions or more than two proportions or between two events or two attributes.

The formula for the *chi-square test* is:

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Where O_{ij} = Observed frequency of cell in row i & column j .

And $E_{ij} = \frac{R_i C_j}{N}$

Where E_{ij} = Expected frequency for cell in row i , column j .

R_i = Sum of frequencies in row i .

C_j = Sum of frequencies in column j .

N = Sum of frequencies for all cells.



As the data collected by the researcher was largely a qualitative one, *chi-square test* was used to find out if there was any significant association between different variables.

3. *t - test*: This parametric test allows the researcher to investigate the effects of one independent variable upon two samples of people. The mean scores of each group are used to calculate ' *t* '. The t-test is the most elementary method for comparing two groups mean scores. Variations of t -test are available for testing independent groups, related groups, and cases in which the population mean is either known or unknown. Therefore t-test is used for testing statistical difference between means of two quantitatively measured independent variables.

The t - test assumes that the variables in the populations from which the samples are drawn are normally distributed. The test also assumes that the data have homogeneity of variance that is; they deviate equally from the mean.

In the journalists' survey questionnaire, as the question regarding experience was left open, the mean could be calculated which lead to the calculation of t value.

The formula for the *t* - test is:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

Where \bar{X}_1 = The *Mean* for *Group1*.

\bar{X}_2 = The *Mean* for *Group2*.

$S_{\bar{X}_1 - \bar{X}_2}$ = The *Standard Error* for the *Groups*.

The *Standard Error* is computed using the following formula:

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\left(\frac{SS_1 + SS_2}{n_1 + n_2 - 2} \right) \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

Where, SS_1 = the *sum of squares* for *Group1*.

SS_2 = the *sum of squares* for *Group2*.

n_1 = the *sample size* for *Group1*.

n_2 = the *sample size* for *Group2*.

4. ANOVA: *Analysis of Variance* is essentially an extension of the t-test. It is advanced in the sense; it can be used to simultaneously investigate several independent variables, also called factors. ANOVA tests for significant differences between two or more group means. It breaks down the total variability in a set of

data into its different sources of variation; that is it “explains” the sources of variance in a set of scores on one or more independent variables.

The ANOVA model assumes that each sample is normally distributed and variance for each group are equal. It also assumes that, the subjects are randomly selected from the population, and the scores are statistically independent, i.e. they have no concomitant relationship with any other variable or score.

The formula for ANOVA is:

$$F = \frac{MS_b}{MS_w}$$

Where, $MS_b = \text{Mean Squares between Groups.}$

$MS_w = \text{Mean Squares within Groups.}$

To calculate *Mean Squares*, the *sums of squares are* divided by the respective *degrees of freedom.*

The *sums of squares (SS)* for *total*, *within* and *between Groups* are calculated using the following formulae.

$$\text{Total SS } (T_{SS}) = \sum X^2 - \frac{(\sum X)^2}{N}$$

$$\text{Within SS } (W_{SS}) = \sum X^2 - \frac{\sum (\sum X)^2}{N}$$

$$\text{Between SS } (B_{SS}) = T_{SS} - W_{SS}$$