

## Research Note

*This department is devoted to research reports, notes, conference résumés and other significant research events. Two copies of write-ups, not exceeding 3,000 words, may be mailed to the Section Editor, Professor David A. Fennell, Brock University, St. Catharines, Ontario, Canada L2S 3A1; e-mail: dfennell@arnie.pec.brocku.ca. In all cases, one copy of the report should be mailed to the Chief Editor, TRR.*

# Use of Automated Content Analysis Techniques for Event Image Assessment

## Noel Scott

Lecturer, School of Tourism and Leisure Management, University of Queensland. e-mail: noel.scott@uq.edu.au

## Andrew E. Smith

Senior Research Officer, ARC Key Centre for Human Factors and Applied Cognitive Psychology, The University of Queensland, Queensland, Australia 4072. e-mail: asmith@humanfactors.uq.edu.au

## Introduction

Schoolies Week is a major annual tourism event in Australia involving students who have completed Year 12, the final year at high school. Similar event activity has been identified in the United States where it is called 'graduation week' (USA Student Travel 2002). In Australia, this event involves Schoolies, visiting the Gold Coast and staying at Surfers Paradise, primarily in commercial accommodation, for around a week each November. Most Schoolies purchase commercial holiday packages offered by a few specialized accommodation brokers. Around 30,000 of these packages were sold in 2004. Some 15,000 of these packages were sold to Queensland Schoolies who primarily attend in the first week and a similar number sold to interstate Schoolies who attend a second Schoolies Week period.

Schoolies Week has a contested image and has been the source of extensive media coverage and debate. Schoolies Week developed as a spontaneous event around 1974 and commercial packages were first offered in 1989. A review of local and State newspaper articles from 1950 to 2004 finds increasing numbers of reports of Schoolies Week from 1988. Early articles in 1988 and 1991 from the *Gold Coast Bulletin*, the major regional daily paper, warned Schoolies to avoid intoxication and promiscuous sex. In 1995, there were reports in *The Courier Mail*, a major Queensland daily newspaper, of a street brawl involving 300 youths in Surfers Paradise during Schoolies Week. In 1997, newspapers reported initiatives by the local police and the Gold Coast City Council to better manage Schoolies Week. In part these initiatives were the result of negative perceptions in tourist

source markets of the Gold Coast as 'attracting undesirables' with potential visitors having 'safety concerns' (Queensland Tourist and Travel Corporation 1997:6)

**Table 1. Selected Headlines Concerning Schoolies Week 1988-1997**

Year	Newspaper Headline	Source
1988	Schoolies are warned not to break the law in relation to underage drinking and vandalism	Dudic, The Gold Coast Bulletin 2/11/88
1991	Schoolies warned on health, wealth hazards scare tactic for young girls participating in promiscuous sex.	Porter, The Gold Coast Bulletin 13/11/91
1995	Schoolies onslaught alert: Wild parties, teenage sex, street brawl and hooliganism	Courier Mail, 18/11/1995,
1997	Planning under way for Schoolies Week invasion: Sgt. Grummitt says we have to turn this dysfunctional period into something positive	The Gold Coast Bulletin, 16/6/1997, p. 6

In 2003 the State Government undertook to manage the Schoolies Week event for two years. The State Government enhanced an existing a program of diversionary activities targeted at Schoolies, enhanced policing of public areas, provided additional support services for Schoolies found to be intoxicated and introduced an education program in schools prior to the event. The study reported in this paper

and derived from publicly available reports examines the change in the content of newspaper articles related to Gold Coast Schoolies Week over the past six years.

The results reported in this paper provide an illustration of the use of automated content analysis software in the study of an event (Schoolies Week on the Gold Coast). The paper also illustrates the use of content analysis to examine the change in the media portrayal (image) of Schoolies Week over the study period. While prior studies of Schoolies Week have been reported, these have been from a risk management and behavioural perspective (Gillespie *et al.* 1991; Ballard *et al.* 1998; Winchester *et al.* 1999; Jansen 2000) or to quantify the economic impact of Schoolies Week (Faulkner 1999; Raybould and Scott 2001). Instead, this study used content analysis to examine change in the textural characteristics of newspaper reports on Schoolies Week in order to examine changes in the public representation of the event.

### Content Analysis

Content analysis has been used since at least the 1950s as a way of analyzing texts (Berelson 1952). Content analysis is a systematic, replicable technique for compressing many words of text into fewer content categories based on explicit rules of coding (Krippendorff 2004). By using content analysis researchers are able to examine large volumes of textual data with relative ease in a systematic fashion to derive quantitative estimates of relationships or patterns. It is a useful technique for discovery and description of focal issues for individuals, groups, institutions, or a society.

The central problems in using content analysis occur in the data reduction process where sentences or other textural units of analysis are classified into categories (Weber 1990:15). These problems concern the validity and reliability of the content analysis process used. The problem of validity concerns the source of the categories used (*a priori*) or produced (*ad hoc*) in the data reduction process as well as the nature of the correspondence between the text and the category. In content analysis studies one approach is to develop an *a priori* code frame and then use this on a sample of the text (say 10 per cent of the material available) to develop a final code frame.

The second validity issue of the correspondence between text and categories is a type of identification problem. Identification is the process of assigning particular examples of a group of items to a category. Two types of identification have been used in content analysis, manifest and latent (Babbie 1992; Fennell 2001). Manifest content analysis uses the surface content of a communication allowing text to be

more coded. Latent content analysis, involves an analysis of the underlying meaning or symbolism of a communication, with the advantage of being more valid.

The issue of reliability in content analysis coding is addressed by the use of software programs that help in analysis of textual data (Alexa and Zuell 2000). There are several reasons for using an automated system for content analysis of text. In many human coded content analysis studies, significant cost and effort is expended to improve inter-coder reliability. Coding errors often occur due to the reliance on human judgement and fatigue (Malloy and Fennell 1998). Human making decisions are also potentially subject to influences which they are unable to report. Efforts to improve the situation may include use of three independent coders (Echtner 2002), which allows inter-coder reliability estimates to be developed (Fennell 2001). Software programs have none of these problems of bias or fatigue and so reliability may be expected to be improved, at the same time reducing cost.

These software programs may also help address the problems of identification and hence improve validity. This is done by automating the coding of the text into categories using a dictionary that is set by the user. Thus, a piece of text will be coded into a category if it contains text that is in the dictionary. The use of a dictionary rather than a single word is important as authors may use synonyms for stylistic reasons throughout a document (Weber 1990). Another reason for using dictionaries is that some words may have multiple meanings. Some software packages are able to incorporate artificial intelligence systems that can differentiate between the same words used with two different meanings based on context.

Much of tourism marketing and management involves the communication of ideas and images of place or type of holiday in textual form. The prime example of such communication is the tourism brochure and these have been subject to content analysis in the area of health tourism (Goodrich and Goodrich 1987) for a mining area (Chon and Evans 1989) and for third world countries (Echtner 2002). More recently, the internet has become a popular media for tourism communication and content analysis has been used to examine websites (Timothy and Groves 2001). Kemp and Dwyer (2003) have used content analysis to examine company mission statements. Media reports have also been discussed as having an impact on tourism, especially in terms of destination image (Stabler 1988; Jenkins 1999) and through reporting of crises and disasters (Faulkner and Vikulov 2001). However, there has been little use of content analysis in tourism studies related to media reporting.

Additionally, in the review of the tourism literature conducted for this study, no use of automated software for content analysis was found. The content analysis software was used in this content analysis study to identify, categorize, and quantify terms, phrases, and expressions in the text that represent Schoolies Week and favourable and unfavourable expressions associated with it.

## Method

The method used here involved several steps including selecting news sources, downloading text, filtering out extraneous paragraphs, coding text, and testing validity (Weber 1990:21-24). The filtering and coding steps were facilitated by use of a content analysis software program called Leximancer Version 2.0 (Smith 2003). The texts of newspaper articles were obtained by performing a keyword search ('Schoolies') from Australian newspaper holdings on the Factiva text base for each year from 1998 to 2004. This database covers all newspapers in Australia including major dailies such as the Courier Mail, Age, Herald Sun and the regional dailies such as the Gold Coast Bulletin. Schoolies Week is a topic of great interest to the local, regional and national press and thus was expected to be well reported. This proved to be the case with increasing numbers of articles found between 1998 (63 articles found) and 2004 (680 articles found). In order to check that the search collected in scope articles, a second independent study was conducted by an independent organization for 2004. This resulted in 688 newspaper articles being identified indicating the scope of the search was adequate.

**Table 2. Number of Newspaper Articles Containing the Keyword 'Schoolies' 1998-2004**

1998	1999	2000	2001	2002	2003	2004
63	89	147	312	573	661	680

(Source: Factiva)

In scope articles were downloaded as plain text and stored in electronic form for analysis. The next step was to 'filter' the text to remove extraneous text using the software application. Significant contamination of the text was observed from two sources. Firstly, a number of Schoolies Week related reports from other destinations apart from the Gold Coast were found. Secondly, the text articles downloaded were found to contain movie reviews of a film about Schoolies Week on the Gold Coast that was released in 2002. A number of sundry other topics of limited relevance to Schoolies' Week were also removed.

These contaminants were filtered out using a Leximancer software technique that involved creation of

categories of related and unrelated words. For example, a category of other Schoolies destination names was created, and used to suppress any article which matched one of the names. Similarly, a category of terms relating to the movie was used to suppress those articles. Finally, an additional function of Leximancer was used that allows lexical concept development. Here words closely related to a particular word or words are found from an initial analysis of the text and stored in a concept thesaurus. In this study, a concept was seeded for Schoolies and Schoolies Week and following thesaurus adaptation, was required to be present in any text segment or the segment was ignored for further coding.

The content analysis software was then used to automate several aspects of the content analysis of this material. For this task, two key Leximancer functions were used: automated dictionary adaptation and automated coding. The categories which were measured in this analysis were as follows: two tonal dimensions – favourable and unfavourable; and six negative factors – alcohol, damage, drugs, sex, 'Toolies', violence. (Note: Toolies is a name given by the media to older men who frequent Schoolies Week, with the intent of preying on Schoolies).

The tonal dimensions were seeded with strongly favourable and unfavourable indicator words using a stock Leximancer list. Seed words for the negative factors were selected from the lists of frequent words from the data. All these seeded categories were then adapted and augmented from the text data by the Leximancer thesaurus builder. The resulting thesaurus was used by Leximancer to code each two-sentence segment of the text with one or more of the tonal and factor categories. Validation by sampled inspection of the coded text segments was then performed, and some iterative refinement of seed words was performed. Learning was performed on 3 sentence segments and classification was performed with 2 sentence segments. These results were further validated by examination of the content of a sample of articles on Schoolies Week by the researchers. This review sought to confirm that the unfavourable words were in fact used in a negative way.

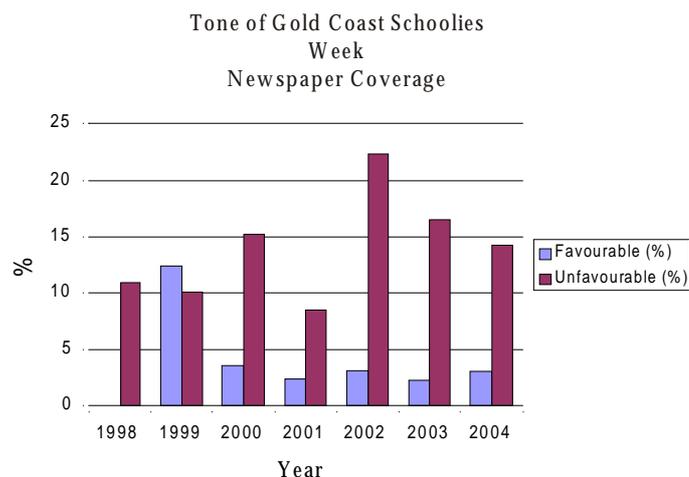
## Results

Final statistics were obtained after coding to show the percentage of relevant two-sentence text segments from each year 1998-2004 which contained each of the tonal and factor codes.

The results of this analysis are shown in Figures 1 and 2. In Figure 1, the percentage of relevant two-sentence text segments from each year containing unfavourable coded words is seen to peak in 2002 and thereafter to decline.

However, the percentage of segments containing favourable coded words has not increased. In Figure 2, the segments are also analyzed to determine percentages that contain six negative factors – alcohol, damage, drugs, sex, Toolies, violence. The results show the association in each year between Schoolie Week and these six negative factors varies. The association with alcohol appears to be declining but the association with drugs and Toolies increased in 2004.

**Figure 1**



**Figure 2**

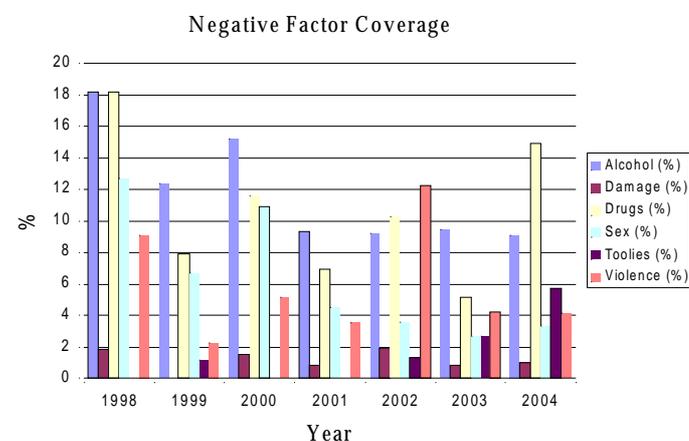


Table 3 provides a selection of headlines from Queensland papers relating to Schoolies Week taken from the study period. The headlines contain numerous references to the potential for danger at Schoolies Week relating to drinking, predators, sexual attacks and drink spiking. These headlines refer to dangers that may occur at Schoolies Week and illustrate the primarily unfavourable tone of newspaper articles found from a review of newspaper article text. This provides some confirmation that the unfavourable words were expressed in a negative manner in the stories used in this research.

**Table 3. Extracts from Media Articles Related to Schoolies Week**

Courier Mail Tuesday 30/11/ 2004	The 'fun police' will be on patrol at Schoolies Week next month as hotel managers enforce a list of hardline rules.
Sunday Mail Sunday 7/11/ 2004	The Schoolies who drink 'til they drop
Gold Coast Bulletin Monday 8/11/2004	Schoolies planning to sneak into pubs and clubs outside Surfers Paradise using fake IDs risk being caught by a new three member taskforce
Gold Coast Bulletin Thursday 25/11/2004	It's danger time for tired revellers as the predators make their play
Gold Coast Sun Wednesday 13/10/2004	Young Gold Coast women are being urged to take extra precautions against sexual attacks in the lead-up Schoolies Week
Courier Mail Friday 26/ 11/2004	Youth workers have warned about 'prank spiking' attacks at the Schoolies Festival by predators seeking to drug teens as a cruel practical joke
Gold Coast Bulletin Tuesday 30/11/2004	This year's Schoolies have been praised for their behaviour, with fewer arrests than last year. But excessive drinking continues to cause concern. Surfers reporter Ryan Ellem looks at the issue

**Conclusion**

The use of automated content analysis software has allowed the quantification of the change in image of Schoolies Week over a six year period. Content analysis using Leximancer software provided a useful method in this study contributing to improved reliability and validity in coding. The results reflect the predominantly negative tone and content of newspaper articles concerning Schoolies Week on the Gold Coast. While some improvement in unfavourable tone over the study period was noted, favourable tone had not increased. This remains a problematic issue for the Gold Coast tourism sector.

Further research using automated content analysis of events and tourism destinations are recommended as it provides a quick and quantitative method for assessing large quantities of text. However, such analyses should be undertaken with care as there are, as has been found in this case, possibilities of contamination of the text data that requires human intervention to eliminate. Also in the case examine here, the identification of articles on the topic of interest was straightforward and the tone and words used not disguised in any way. In fact the articles use terms (sex, danger, drugs, etc.) that 'jump out' of the text. In cases were the text used is more subtle or nuanced the use of software such as Leximancer may require further testing.

## References

- ALEXA, M. and ZUELL, C. (2000). Text Analysis Software: Commonalities, Differences and Limitations: The Results of a Review. *Quality and Quantity* 34: 299-321.
- BABBIE, E. (1992). *The Practice of Social Research* 6th edition. Belmont, CA. Wadsworth.
- BALLARD, R., CURD, D. and ROCHE, A. (1998). The Meaning, Madness and the Magic of 'Schoolies' Festival. Paper presented at the APSAD National Conference. Sydney.
- BERELSON, B. (1952). *Content Analysis in Communication Research*. Glencoe, Ill. Free Press.
- CHON, K. and EVANS, M. (1989). Tourism in a Rural Area - A Coal Mining-county Experience. *Tourism Management* 10(4): 315-321.
- ECHTNER, C. M. (2002). The Content of Third World Advertising: A 4A Approach. *International Journal of Tourism Research* 4: 413-434.
- FAULKNER, B. (1999). *Economic Impact of 'Schoolies' on the Gold Coast Economy*. Gold Coast. Centre for Tourism and Hotel Management.
- FAULKNER, B. and VIKULOV, L. (2001). Katherine, Washed Out One Day, Back on Track The Next: A Post-mortem of a Tourism Disaster. *Tourism Management* 22(4): 331-344.
- FENNELL, D. A. (2001). A Content Analysis of Ecotourism Definitions. *Current Issues in Tourism* 4(5): 403-421.
- GILLESPIE, A., DAVEY, J., SHEEHAN, M. and STEADSON, D. (1991). Thrills without Spills: The Educational Implications of Research into Adolescent Binge Drinking for a School Based Intervention. *Drug Education Journal of Australia* 5(2): 121-127.
- GOODRICH, J. N. and GOODRICH, G. E. (1987). Health-Care Tourism: An Exploratory Study. *Tourism Management* 8(3): 217-222.
- JANSEN, E. J. (2000). *Do Service Providers of 'Schoolies' at the Gold Coast identify 'Schoolies' as a Form of Event Tourism?* Unpublished Bachelor of Business -Travel and Management (Honours) degree. University of Queensland. Ipswich.
- JENKINS, O. (1999). Understanding and Measuring Tourist Destination Images. *International Journal of Tourism Research* 1(1): 1-15.
- KEMP, S. and DWYER, L. (2003). Mission Statements of International Airlines: A Content Analysis. *Tourism Management* 24: 635-653.
- KRIPPENDORF, K. (2004). *Content Analysis : An Introduction to its Methodology* 2nd edition. Thousand Oaks, CA. Sage.
- MALLOY, D. and FENNELL, D. A. (1998). Codes of Ethics and Tourism: An Exploratory Content Analysis. *Tourism Management* 19(5): 453-461.
- QUEENSLAND TOURIST AND TRAVEL CORPORATION (1997). *Gold Coast Domestic Leisure Marketing Strategy 1997-1999*. Brisbane. Queensland Tourist and Travel Corporation.
- RAYBOULD, M. and SCOTT, N. (2001). *'Schoolies' Week Gold Coast 2001: Report of the Results of a Visitor Survey and Estimation of Economic Impacts*. Gold Coast. Cooperative Research Centre for Sustainable Tourism.
- SMITH, A. E. (2003). Automatic Extraction of Semantic Networks from Text using Leximancer. Paper presented at the HLT-NAACL 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics - Edmonton. Alberta, Canada.
- STABLER, M. J. (1988). The Image of Destination Regions: Theoretical and Empirical Aspects. In Goodall, B. and Ashworth, G. (Eds.) *Marketing in the Tourism Industry: The Promotion of Destination Regions*. London. Routledge: 133-159.
- TIMOTHY, D. and GROVES, D. (2001). Research Note: Webcam Images as Potential Data Sources for Tourism Research. *Tourism Geographies* 3(4): 394-404.
- USA STUDENT TRAVEL (2002). *Graduation Week*. [http://www.usastudenttravel.com/usa/intro\\_gw2003.htm](http://www.usastudenttravel.com/usa/intro_gw2003.htm). Accessed on 7 June 2002.
- WEBER, R. P. (1990). *Basic Content Analysis* Vol. 2nd. Newbury Park, CA. Sage.
- WINCHESTER, H. P. M., MCGUIRK, P. M. and EVERETT, K. (1999). Celebrations and Control: Schoolies Week on the Gold Coast Queensland. In Teather, E. (Ed.) *Embodied Geographies: Spaces, Bodies and Rites of Passage*. London. Routledge.

Submitted: April 8, 2005

Accepted: April 28, 2005