

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Patterns of Viewing Behaviour during Advertisement Breaks in Different Television Programmes

A thesis presented in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Marketing at Massey University

Maureen Syn

2002

MASSEY UNIVERSITY

APPLICATION FOR APPROVAL OF REQUEST TO EMBARGO A THESIS
(Pursuant to AC98/168 (Revised 2), Approved by Academic Board 16.02.99)

Name of Candidate: Maureen Sun I.D. Number: 86029022

Degree: Ph.D Dept / Institute / School: Marketing, Business

Thesis Title: Patterns of viewing behaviour during advertisement
breaks in different television programmes

Name of Chief Supervisor: Dr. M. Brennan Telephone Ext: 5576

As author of the above named thesis, I request that my thesis be embargoed from public access until (date) 1 February 2004 for the following reasons:

- Thesis contains commercially sensitive information.
- Thesis contains information which is personal or private and / or which was given on the basis that it not be disclosed.
- Immediate disclosure of thesis contents would not allow the author a reasonable opportunity to publish all or part of the thesis.
- Other (specify) : _____

Please explain here why you think this request is justified:

Additional time is needed for possible publication of parts
of the thesis.

Signed (Candidate): Maureen Sun Date: 30.1.03

Endorsed (Chief Supervisor): M Brennan Date: 31.1.03

Approved / ~~Not Approved~~ (Representative of VC): M Blogg Date: 12/2/2003

Note: Copies of this form, once approved by the representative of the Vice Chancellor, must be bound into every copy of the thesis.

ACKNOWLEDGEMENTS

I wish to thank my supervisors Mike Brennan and Zane Kearns for the guidance and kind support they have given to me. Many thanks to Mike who constantly advised me on the next step and, gave me ideas, suggestions and feedback to help me move along. My sincere thanks to Zane who was always there to explain and clarify issues for me.

Acknowledgement should also be given to the Department of Marketing at Massey University and the International Pacific College who funded the equipment used for the research. Without the financial assistance, it would take many more years to initiate and to finish this project.

I owe a huge debt to my husband, Peter Chow who has consistently over the years supported me financially and emotionally. He was also the "labourer" who had to transport and physically carry the recording equipment used to collect the data for this research. But most of all, his encouragement, pragmatism and excellent sense of humour had helped me enormously and enabled me to get over many stressful situations during the stages of this research.

ABSTRACT

This thesis sets out to examine the viewing behaviour of television audiences during advertising breaks. The results reported are based on the analysis of in-home video recordings of 14 families watching television in their own homes over an 8-day period.

In total, the viewing of 6017 advertisements was examined, and the status of the viewer was recorded at 3-second intervals. Each 3-second observation was accorded one of four possible states: the viewers were directing their attention to the screen (*EOS only*); they were directing their attention to the screen but engaging in additional behaviours (*EOS + behaviour*); they were not directing their attention to the screen (*No EOS*); or they were not present.

Across all advertisements, visual attention to the screen was generally low, despite viewers being present for a high proportion of the time during the ad breaks. Viewers were present during the ad breaks for 78% of the time, but were only paying attention to the screen fully 14% (*EOS only*), and looking as well as doing something else simultaneously 11% of the time (*EOS + behaviour*). About half the time during the ad break (52%), no one was looking at the screen.

However, more significantly for advertisers, ad-viewing behaviour differed by programme type. For example, some programmes could command as much as eight times the proportion of eyes-on-screen time than other programmes. Moreover, the proportions of particular ad-viewing behaviour differ in an unpredictable way for different programmes. Thus different conclusions might be reached when for an example, an advertiser sought to optimise a media plan in terms of *EOS only* than when the choice was made to optimise another characteristic of ad viewing, say eye-on-screen and doing something else simultaneously (*EOS + behavior*).

In examining three specific methods of classifying programmes, none of them provided workable descriptions of ad-viewing behaviour. This problem manifests itself in a number of ways, but primarily, the aggregation of programmes into categories masks considerable within-group variation, pointing to the need to consider the effect of programmes at a much more disaggregate level, perhaps even to the extent of considering each episode or segment of the programme before the ad break separately.

Thus, this research has raised questions about programme ratings and demonstrated that the size of an audience may not reflect how attentive that audience is. Even programmes of small audience sizes may have a high proportion of attentive ad viewers and may thus be of interest to advertisers since advertisements placed within such programmes may be equally (if not more) effective than programmes with a larger audience size.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	i
ABSTRACT	ii
TABLE OF CONTENTS	iv
LIST OF APPENDICES	viii
LIST OF TABLES	ix
1. INTRODUCTION	1
1.1 Introduction	1
1.2 Objectives of Study	8
1.3 Outline of the Chapters	9
2. LITERATURE REVIEW	10
2.1 Introduction	10
2.2 Television Ratings Measurement	12
2.2.1 Diary method	12
2.2.1.1 Non-meter diaries	12
2.2.1.2 Metered diaries	14
2.2.2 Peplemeter research	17
2.2.2.1 Historical overview	17
2.2.2.2 The Peplemeter technology	18
2.2.2.3 Peplemeter use in New Zealand	20
2.2.2.4 Criticisms of Peplemeter	21
2.3 Television Audience Viewing Research	33
2.3.1 Introduction	33
2.3.2 Peplemeter studies	34

2.3.2.1	Introduction	34
2.3.2.2	Ads	34
2.3.2.3	Programmes	38
2.3.2.4	Summary	39
2.3.3	Observational research	40
2.3.3.1	Historical overview	41
2.3.3.2	Viewing patterns	43
2.3.3.3	Relationship between programme and ad viewing	47
2.3.3.4	Summary	51
2.3.4	Effects of programme impact on viewing behaviour	52
2.3.4.1	Introduction	52
2.3.4.2	Demandingness according to Barwise and Ehrenberg (1988)	53
2.3.4.3	Involvement	56
2.3.4.4	Summary	61
2.4	Summary and Conclusions	63
3.	RESEARCH METHOD	66
3.1	Introduction	66
3.2	Sample	67
3.3	Equipment	72
3.4	Programme Involvement Rating Scale	74
3.5	Procedure	75
3.6	Coding	77
3.6.1	Coding overall viewing behaviours	77
3.6.2	Ad pod selection	79
3.6.3	Behaviour during ad breaks	80
3.6.4	Programme codes	81
3.7	Analysis	85

3.7.1	Complete ad pods	85
4.	RESULTS	87
4.1	Introduction	87
4.2	Overall Viewing Patterns of the Participants	88
4.2.1	Viewing time	88
4.2.2	Types of programmes	90
4.3	Behaviour during Advertisement Breaks	92
4.4	The Relationship between Programme Involvement and Ad Involvement	95
4.5	Ad-break Behaviours by Programme Type Using the Classifications Developed in Barwise and Ehrenberg (1988)	96
4.5.1	Comparison of behaviour during <i>demanding</i> and <i>less-demanding</i> programmes	97
4.5.2	Comparison of programmes within <i>demanding</i> programmes	98
4.5.3	Comparison of programmes within <i>less-demanding</i> programmes	99
4.5.4	Comparison of visual attention levels during <i>demanding</i> and <i>less-demanding</i> programmes	100
4.5.5	Comparing viewing behaviour within <i>demanding</i> programmes	101
4.5.6	Comparing viewing behaviour within <i>less-demanding</i> programmes	102
4.5.7	Summary	103
4.6	Ad-break Behaviour by Genres	105
4.6.1	Quality of viewing across different genres	107
4.6.2	Summary	114
4.7	Viewing Behaviours by Impact Categories Developed by Hoffman and Batra (1991)	116

4.7.1	Viewing behaviours associated with programme impact scores	117
4.7.2	Visual attention during presence	119
4.7.3	Comparing results on impact categories	121
4.7.4	Relationship between programme types by genres and impact categories	125
4.7.5	Summary	127
5.	DISCUSSION	129
5.1	Introduction	129
5.2	Presence Versus Attention	134
5.3	Viewing Behaviour and Programme Classification Systems	137
5.3.1	Barwise and Ehrenberg (1988) classification	138
5.3.2	Genre classification	140
5.3.3	Hoffman and Batra (1991) classification	142
5.4	Conclusion and Contribution	146
5.5	Limitations and Implications for Future Research	150
5.5.1	Limitations of research	150
5.5.2	Implications for future research	151
6.	REFERENCES	155
7.	APPENDICES	165
	Glossary	181

LIST OF APPENDICES

Appendix A:	The C-Box	165
Appendix B:	Participant's Programme Evaluation Questionnaire	166
Appendix C:	Participant's Information Sheet	169
Appendix D:	Participant's Consent Form	172
Appendix E:	Summary of Participant Families	173
Appendix F:	Rationale for 30-second Criteria for Ad Pod Selection	174
Appendix G:	Examples of Ad-viewing Behaviour	179

LIST OF TABLES

Table 1	ADRATIO by programme type	36
Table 2	Summary details of participants	71
Table 3	Household viewing time	88
Table 4	Time individual household members were present	89
Table 5	Proportion of time viewers were present during different programmes	90
Table 6	Viewing behaviour during complete ad pods	93
Table 7	Attention during complete ad pods	94
Table 8	Ad-break behaviours by programme type using Barwise and Ehrenberg (1988) categories	98
Table 9	Ratios of visual attention to actual presence	100
Table 10	Viewing behaviour by genre categories	106
Table 11	Visual attention and presence	108
Table 12	Viewing behaviour during comedy programmes	113
Table 13	Viewing behaviour by impact scores	117
Table 14	Viewing behaviour by impact scores during time present	120
Table 15	Viewing behaviour by Hoffman and Batra (1991)	123
Table 16	Viewing behaviour of this study	123
Table 17	Impact scores by programme types	126
Table A1	Comparison of ad-viewing behaviour	176
Table A2	Comparison of ad-viewing behaviour using data from viewers' personal evaluation of programmes (Hoffman & Batra, 1991)	177

CHAPTER ONE

INTRODUCTION

1.1 Background

Billions of dollars are spent annually on television advertising. In the United States, the 2001 television expenditure was projected as over US\$50 billion (Coen, 2001). In the UK, total television adspend for the year 2000 amounted to about £3,949m (World Advertising Research Centre, 2002). As Asia's fastest growing television market, China spent about US\$7 billion on TV advertising in 2000 out of its total advertising budget of US\$11 billion (Green, 2001). New Zealand advertisers paid a total of around NZ\$479m for television advertisements in 2001 (Admedia, 2002).

Individual advertisements can also be costly. For example, the most expensive advertising space in 2000, was sold for US\$2.2 million for a 30-second commercial during the screening of SuperBowl games in January (Lealand & Martin, 2001). In New Zealand, the cost of placing an ad on the main national channel, TV One, is at least \$8,500 for a 30-second ad during primetime (6 pm-10.30 pm). Advertisements rates vary, depending on the popularity of the programme in which they are to be placed and the time of screening. However, regardless of how much advertisers are charged for their advertisements on television, the investment in television advertising is unequivocally substantial.

This worldwide multi-billion-dollar expenditure on television advertising underlines the importance given to this part of the marketing mix. It also indicates the reasons why advertisers want to know whether the advertising used is effective and justifies the considerable investment they have made. For any particular advertisement to be successful, it must undergo the complex

process of being screened, attended to and processed by the consumer to have some ultimate effect. As a consequence, an important television advertising decision concerns when to screen advertisements so as to obtain the best results. Traditionally, these decisions have been based on TV audience programme viewing ratings.

The standard technique currently used worldwide to obtain television audience viewing ratings is the Peoplemeter. With this device, data on television viewing are collected electronically, based on viewers' self-reports of their television viewing. Viewers in the Peoplemeter panel "log in" and "log out" by pressing buttons on a remote control linked to the television set as they enter or leave the viewing room. This measurement gives an indication of the number of people present in the room as the programme is being screened.

There are, however, several major criticisms of Peoplemeter data. One concern is that Peoplemeter data cannot report on the actual number of attentive viewers during the screening of the advertisements. The Peoplemeter does not provide data on the quality of ad viewing – an aspect which many advertisers are interested to know. Another criticism is that the data may be inaccurate because of the limitations of the methodology. A major concern is that the technology relies entirely on participants to press a button to indicate their viewing, and this introduces the possibility of human error. But perhaps the most important concern is that because the Peoplemeter provides measures of programme viewing, this might overstate the size of audiences viewing the advertisements within any given programme.

It is widely acknowledged that people often avoid advertisements either physically by leaving the room or mechanically by channel flicking. According to some media commentators (e.g., Moriarty & Everett, 1994; Gunter, Furnham & Beeson, 1997), this trend will become more significant as modern life becomes more and more technologically complex with the advent of satellite

television and digital media technology. This has negative implications for advertisers. The huge investment advertisers put into their television advertisements would be wasted if opportunities to see their advertisements were greatly reduced. A key issue, then, is whether the Peoplemeter provides an accurate estimate of ad avoidance.

So far, the evidence of ad avoidance is conflicting. Abernethy (1991) reported an overall 32% commercial avoidance norm, and Clancey (1994) found viewers avoid commercials 31% of the time. By contrast, Danaher (1995) reported that 90% of the audience watched commercials in New Zealand. Danaher's (1995) study investigated the behaviours of participants for a week and was based on data collected via the Peoplemeter.

Similar findings have been reported by Lynch and Stipp (1999). Based on data obtained from Nielsen Media Research in the US and the SMART laboratories in Philadelphia (via the Peoplemeter system), they reported that on average, over 90% of viewers stayed for the commercials during primetime programmes, and there was evidence that ad avoidance was even less frequent during the highest rated television programmes. These conflicting estimates of ad avoidance are of concern to TV advertisers.

However, even if ad avoidance is not a major problem, the Peoplemeter system, at best, simply records presence in the room, and gives no indication of the range of viewing activities that take place when the advertisements are screened. Even when people are present in the TV room, the quality of the viewing behaviour may vary enormously, ranging from attentive viewing or engaging in complementary activities (such as eating or doing household chores or hobby crafts), through to competitive activities (such as reading, talking or sleeping) that directly compete with TV viewing. It is thus quite possible that the gross audience size suggested by ratings figures is far greater than the number of viewers who actually attend to any particular advertisement. This

raises the question whether or not the measure of "presence" provided by Peoplemeter provides a valid surrogate measure of "attention" to television advertising.

As Peoplemeter technology cannot measure "attention", a number of studies have specifically used measures of visual attention to investigate viewers' involvement with programmes and ads (Anderson, 1986; Krugman, Cameron and White, 1995). One commonly used measure is *eyes-on-screen* (EOS): the measurement of how long individuals looked at the screen. To determine the amount of time viewers were actually looking at the screen, trained observers timed and recorded the duration of each continuous eyes-on-screen event and the number of such events (Anderson, 1986; Krugman, Cameron & White, 1995). Although this method is not without limitations, according to Thorson (1994), visual orientation to the screen is highly correlated with the internal processes of the viewers' minds and provides a valid measure of involvement.

While viewer behaviour is an important consideration, the viewing of advertisements does not happen in isolation; it is always within the context of programmes. Different programmes have different impacts on viewers, and this may in turn influence the amount of attention viewers pay to the advertisements. Thus from an advertiser's perspective, it will be important to know what effects programme characteristics have on viewers' attention to the advertisements. In this way, it would be possible to identify the types of programmes where advertisements are most likely to be effective. To know the relationship between programmes and advertisements, one needs first to be able to differentiate between types of programmes. In studies of television viewing, different approaches to classifying programme types have generally been used.

The most common grouping method is to classify programmes of similar content or programme format characteristics into "genres", such as news,

situation comedies, movies and so on. This type of classification was used by Danaher (1995), and is used by Peoplemeter studies.

Genres are most commonly used in the TV industry. As television audiences increased in size over the years, and programming output increased, there was a need to group and describe similar programmes in terms of their "shared stories, narrative and visual styles, and intention" (Lealand & Martin, 2001, p 60). This also helped the audiences to know what to expect from each type of genre. For example, a half-hour sit-com is expected to be a funny and relaxing type of light entertainment. However, although genre classification could systematically categorise programmes, it is increasingly more difficult to distinguish between genres as television programming changes to cater to differing audience tastes.

There are now new strands of genres that have the characteristics of two or more types of "traditional" genres. Reality TV is cited as an example. According to Lealand and Martin (2001), Reality TV is a new "hybrid" genre that is "a combination of a game show, soap opera and nature documentary, with a delicate balance of predictability (someone will be eliminated) and unpredictability (who will it be?)" (p. 141). Even within each category of genre, there are differences. Different Reality TV programmes, such as "*The Osbournes*", "*Big Brother*" or the "*The Amazing Race*" have different country settings, different themes, and attract different target markets. These differences show how programmes within the same genre category may still be intrinsically different and how difficult it is to assign the diversity of programmes to appropriate or "correct" genre categories. Such difficulties also demonstrate that programme classification is inherently problematic.

In another study about television audiences, Barwise and Ehrenberg (1987, 1988) studied programmes by the way they made "demands" on viewers. For example, according to them, news and documentaries were *demanding*

programmes, while entertainment programmes such as situation comedies were *less-demanding*.

Other researchers measured a programme's "impact" in terms of how much the programme made viewers feel aroused, active, suspenseful or tense. For example, action-adventure programmes were assumed to be more "involving" than situation comedies (Kennedy, 1971; Soldow & Principe, 1981). Similarly, Hoffman and Batra (1991) suggested programmes could be grouped in terms of how viewers felt each programme had involved them via programme-rating scales.

Thus, despite the problems associated with programme classifications, different approaches have been used. In studying the behaviour of ad viewers and the effects of programmes have on viewers, it is also necessary to find out whether different programme classification methods affect viewing behaviour. If they do, this will shed some light on the value of buying air space for TV advertisements during specific programme types rather than during individual programmes.

Some issues about the validity of the information provided for advertisers have been raised. There are doubts about the validity of the present ratings system as one cannot assume that ratings for programmes are similar to ratings for ad breaks. Moreover, ratings do not show the number of attentive viewers actually watching the advertisements. In other words, qualitative measurements are needed to identify the different levels of attention viewers pay to the screen. A common measure used in many "attention" studies is *eyes-on-screen* (EOS), which can be used to investigate the actual attention paid to advertisements during different programmes. This measure would provide data about the types of programme characteristics that involve viewers. Information about how viewers behave in different programme contexts would further assist advertisers to make effective media placement decisions that would maximise the return on

their investment. This thesis aims to cover this gap by providing details to show how viewers pay attention (using EOS as the measurement) and whether their ad-viewing behaviour differs as a result of the programmes that delivered the advertisements.

1.2 Objectives of Study

The general aim of this study is to deepen our understanding of the relationship between television programming and advertising in terms of viewing behaviours during commercial breaks. To achieve the general aim, specific objectives were formulated and are outlined below.

1. To compare the proportion of time viewers spend visually attending and not attending to advertisements during ad breaks.
2. To examine the relationship between different levels of visual attention and presence during ad breaks.
3. To compare ad-break viewing behaviour across different programme classification methods used to group programmes.
4. To examine the relationship between visual attention and viewers' self-rated level of involvement with programmes.

1.3 Outline of the Chapters

Chapter Two reviews the literature on the dominant methods in measuring television audiences, the importance of measuring attention to commercials, the relationship between programme environment and ad viewing, and whether different types of programmes affect the viewing of commercials.

Chapter Three details the methodology used in this study. This chapter contains a detailed description of the equipment used to collect data on viewing behaviour, as well as the procedure used to code and analyse the data.

Chapter Four outlines the research findings. This chapter presents the results in terms of the viewing behaviours displayed by viewers during the advertisement breaks. It presents findings on the attention levels paid to advertisements and the quality of the viewing. The remainder of the chapter describes whether the quality of the viewing differed in different programmes and discusses the validity of three programme classification approaches used to group TV programmes.

Chapter Five discusses how the results compare and contrast with earlier studies on attention to commercials and what the implications for advertisers are. This chapter also points out the limitations of this study and the implications for future research in this area.

CHAPTER TWO

LITERATURE REVIEW

This section introduces and reviews the main methodologies employed in the measurement of television audiences. It discusses first the advantages and criticisms of these methods, and then how "attention" to commercials could be measured and is related to various aspects of the viewing environment, in particular to different types of programmes. The chapter then provides an overview of earlier research on the effects of programme involvement on the receptivity of audiences to commercials. It looks also at some methods used by different researchers to group programmes.

2.1 Introduction

In just about every developed country in the world, television broadcasters, media owners, advertisers and their agents invest large sums of money every year in measuring the size and composition of their audiences. This information is normally provided by media research agencies that attempt to provide an accurate measure of the numbers and type of people who have an "opportunity to see" (known in the industry as OTS) an advertisement placed by an advertiser. "Opportunity to see" is usually measured in terms of the number of people present in the room during the time the television advertisement was being screened. This information may also include the description of the audience, and what they think about the programmes.

In general, audience measurement methods can be either continuous studies or independent studies conducted periodically. Continuous studies recruit a constant stable panel and monitor these homes over a period of time, mainly to

measure viewers being present in the viewing room. Periodic stand-alone studies involve independent samples for each measurement period, and are conducted mainly to acquire a clearer picture of audience attention to programmes and to commercials.

Data on media audiences are normally obtained by using one or more of three main instruments (Kent, 1994) :

- questionnaires,
- diaries,
- electronic recording devices.

While questionnaires may be completed either by the respondents or by the interviewer, diaries rely solely on the respondents. In the diary method, respondents make an entry about their viewing behaviours over a fixed time period. As for electronic recording devices, the earliest types used in the television industry were "TV meters"—devices wired to individual TV sets, which provided information about the status of the television at various time periods.

"TV meters" were at times used in combination with diaries. These two methods (either on their own or in combination) continued to be used in many countries until the middle of the 1980s. After about 1985, a more advanced technology, known as the Peoplemeter, was introduced in Europe. This device recorded television set data electronically and required participants to indicate their viewing by pushing pre-designated buttons on remote control handsets. Since its inception, this methodology has replaced most conventional diaries and has become the dominant audience measurement method adopted in many countries.

The next section traces the development of audience measurement methods until the advent of Peoplemeter type technology.

2.2 Television Ratings Measurement

2.2.1 Diary method

Most continuous audience measurement provides information on the ratings of different programmes. This is the estimated size of its audience expressed as a percentage of the population size. For many years, self-completed diaries were used in the United Kingdom and in the United States as the principal method to record the day-to-day television viewing of adults and to provide information on programme ratings. There were two kinds of diary-based panel households: the continuous meter-controlled diary panel and the non-meter diary panel. In the UK, non-metered diaries were first used in the early 1950s, while metered diaries started after 1956 (Gane, 1994).

In non-meter diary panels, viewers record in the weekly diary which channel they viewed, as well as their presence in the room when the set was switched on. In meter-controlled diary panels, "TV meters" report information about the changes to the status of the television set (i.e. on or off and the channel being viewed), while at the same time, family members record their behaviours in the household's weekly diary.

2.2.1.1 Non-meter diaries

All diaries have the capacity to record behaviour that is repeated at intervals. They may be structured or unstructured. Fully structured or pre-coded diaries contain full listings of all the channels and programmes of a particular area that are provided for each family. Entries by time segment on a daily basis, often 15-minute or 30-minute periods, are pre-printed for the entrant. The respondent may then be asked to tick off columns on the page to indicate the channel and the programme watched. In contrast to structured diaries, unstructured ones

require the respondents to write in details on the channels, stations or programmes as appropriate (Kent, 1994).

There were, however, a number of criticisms of diary-based methodologies. It was found that people usually did not fill in their diaries while they watched television (Rubens, 1978). They often reconstructed their activities (from memory) a couple of days after the viewing or even just before sending off the diaries. The diary might also be deliberately falsified either by "omission of some media use or by the inclusion of imaginary uses" (Kent, 1994). In addition, the diary method suffered from "fatigue" because of the repetitive nature of constantly recording actions (Bennett, 1983). All of these factors may bring about inaccuracies in the reporting.

Moreover, as channel switching became more frequent as a result of the increased number of channels, parts of the individual's viewing also did not enter the processed record. The diary completion exercise had the fundamental problem of having to work in fixed, usually 15-minute periods, and information less than 15 minutes was not entered. Furthermore, only viewing of more than half (eight minutes) of the 15-minute block needed to be recorded. If the diary was correctly filled, an individual watching from 7.55 to 8.05 did not "count". Similarly, someone viewing from 7.30 to 8.02 was not deemed to have seen the 8.01 commercial. Conversely, the viewer watching from 7.00 to 7.10 then switching to another channel was recorded as having seen the 7.14 commercial. This incomplete gathering of viewing information was further aggravated by the wide use of remote control sets, which made channel switching and zapping easier.

There were also doubts about the quality of diary keeping for the other sets in multi-set households. The housewife (in the household) was usually the "main custodian" of the diary and the main person to ensure accurate completion of the diaries. In multi-set homes, the knowledge of the "main custodian" about

other television sets in the house might not be as good as for the main set. The main custodian might not know what was being watched in the other viewing rooms, so could not ensure the diary entries were always accurately entered. It was thought that about 6–7% of the decrease in ratings (during autumn) could be attributed to "this measurement gap" (Bennet, 1983).

Other more general comments about the television audience data gathered by diaries were that they were not comprehensive enough to aid media planning, not specific enough to permit accurate media buying, and not fast enough to track changing patterns of the viewer (Bennet, 1983).

2.2.1.2 Metered diaries

Electronic measurement devices have been used to monitor television audiences for over 50 years (Gill, 2000). TV meters were first introduced in the US as early as the late 1940s. In 1956, the UK adopted this methodology, and became the first country in Europe to use electronic audience measurement devices (Gane, 1994). These TV meters (also known as "set meters") were attached to the TV set in the panel household and automatically recorded minute-by-minute whether the set was switched on and to which channel it was tuned. In addition, panel members were required to record their presence in the room in 15-minute periods in the household's weekly diary manually (Ehrenberg & Twyman, 1965). In the 1960s and 1970s, the combination of the "set meters" and the diary (metered diary method) was used in many parts of the world, and was considered a good measure of viewing behaviour (Buck, 1987).

By the early 1980s, however, changes in the technical television environment also meant changes had to be made if television audience measurement was to maintain its credibility. These changes included the introduction of more channels on TV, the increasing popularity and use of the VCR, multi-set homes,

and out-of-home viewing when portable television sets were used. These changes posed problems for the existing system because the meters were designed to work in a simple single-set home environment, and where the family unit was expected to be watching television together.

By the middle of the 1980s, newer types of meters were installed in the UK. The bulky old meters were replaced by smaller "slave" meters capable of being placed on either static or portable television sets (Buck, 1987). These newer sets transmitted information to a master meter in the home, which then passed the day's data to a centralised computer. The development of these types of meters opened the way for today's more advanced electronic methods.

However, due to the changes in meter technology, the self-completed diary and the meter systems of collecting data became over time more and more incompatible. Rawlings (1983) commented that the meter, as a new technology, was able to measure television set data quite accurately, while the diary was a "rather primitive" method used to record the accompanying audience data. The electronic meter method was also able to record a very high sampling of television "viewing" (e.g., every 20 seconds), while the manual diary method included information only when it was longer than eight minutes of each 15-minute block in the hour. In other words, the meter (providing information on the television set on/off status and channel tuned) was able to show data by 20 second, but the diary method still provided information on individual viewing by 8 minutes. Furthermore, the information provided by the diaries in this format, was of limited use to some advertisers. For example, advertisers interested in a 30-second commercial spot found diary entries (measured by 15-minute blocks) too general and lacking in precision (Buck, 1987).

Together with the introduction of more channels, there was also the growth of the VCR. The VCR could be used to view "live" broadcasts (programmes viewed as they were being screened) and "time-shift" viewing (programmes

viewed at a later time after it was screened). According to Bennet (1983), about 13% of the half-hour drop in individual viewing in the October to December 1982 quarter, was due to the non-reporting of "live" viewing through the VCR tuner. A further 5% was due to time-shift viewing of broadcast programmes (Bennet, 1983). This was largely because the set meters that recorded the changes in the television set had not been designed to process information from the VCR tuner. As a result, all VCR viewing had been "discarded" during the processing. However, in subsequent years, even when the problem was rectified and VCR information could be included in the analysis, the meter technology was still unable to distinguish which channel was being watched during the VCR viewing. Thus while it was possible to detect "time-shift" viewing, this information was not included in the ratings calculation for recorded programmes.

There were also problems with recording the viewing in homes that had more than one television set. In multi-television sets households, about 10% of total viewing, around half an hour per day on average, a second television set was turned on, yet the same diary editing rules as for the main set were used. If there was a mismatch of half an hour per day or two and a half a week, then the viewing records were discarded (Bennet, 1983).

Lind (1984) had three criticisms about the UK metered diary method. The first was that the conversion of data into relevant usable results was always slow. For example in the UK, the results were published weekly in two densely packed volumes, filled with minute-by-minute information on the viewing of four channels and a quarter-hourly viewership of commercial breaks. The second criticism was that the system also did not fully measure the audiences of newer channels such as cable TV, which tended to have smaller audiences. For example, in some cases, the data for such audiences were only based on as few as thirty sets. The third criticism is that BARB (Broadcaster's Audience Research Board) who was responsible for the research, was not prepared to

discuss the results openly, and this caused increased suspicion of the actual value of the results. As the users of these results became more suspicious of the accuracy of the figures, the credibility of the results lowered.

2.2.2 Peoplemeter research

The diary-based audience measurement systems reviewed in the previous section have largely been supplanted by what has become known as Peoplemeter technology. This has come about as a response to the changing and more complex television viewing environment, for example, increases in multi-set homes and channels (Gane, 1994). According to Billet (1985), the diary method was satisfactory when viewing choices were simple and audiences not so fragmented. By the 1980s, as the television environment and technologies became more complex and technologically sophisticated, it became increasingly difficult to use, administer and defend the simple self-completion diary or, indeed, any manual measurement system.

2.2.2.1 Historical overview

The first Peoplemeter system was installed in the UK and in Italy in 1984, and in Ireland, Switzerland and West Germany in 1985 (Gane, 1994). France was perhaps the only country in Europe, which thoroughly debated and tested the accuracy of the Peoplemeter approach. It eventually adopted this system nationally in 1987. By the end of 1995, 48 Peoplemeter systems were operational in 41 countries of the world. By 2000, it was being used in 70 countries (Gill, 2000), including newly developing countries such as the Republic of China and countries in the Eastern Bloc.

2.2.2.2 The Peoplemeter technology

In a Peoplemeter system, the equipment generally comprises a meter fitted on top of the TV known as the meter display unit (MDU), a remote control handset, and a central data storage unit (CDSU) (Sharot, 1994).

The MDU records the station and channel selection and the status of all other viewing equipment such as the VCR or cable, and also displays who is watching the set. The system includes a smart probe, a device link to the household's VCR, which records the channel to which the VCR is tuned and whether the VCR is playing a tape, in pause, search or fast-forward mode. The system can also detect whether the VCR is in playback, as well as all other types of recordings played on the recorder.

Unlike the diary-based methods, where respondents had to fill in a diary of their television viewing, in this method respondents simply press a button on a remote control handset that signals viewers have started watching; when the viewers stop watching, they "log-out" by again pressing a button.

The CDSU is located in an inconspicuous corner of the house and contains a modem located near the telephone point. It may act as the main meter for all other MDUs in the house. It communicates, interrogates and copies all switching records or changes in viewing status into its own memory. The agency's central computer then dials up for the information in the early hours of the morning. The modem is programmed to go on-line at pre-specified times during the night, and the unloading takes between 30 and 60 seconds each time.

During the mid-1980s, on-line results were offered and subscribers dialled in for overnight results. In the 1990s, subscribers could take the entire database on tape and download for processing in their own mainframes and PC systems.

Further developments of the methodology have attempted to remove the need for panellists to "indicate" watching via the remote control handset. Techniques tried by AGB and Nielsen have included image recognition using neural networks. In this system a computer identifies the people in the room and can maintain that image as they move around the room, or for example, change places on the couch. It can even detect faces hidden behind a book or momentarily turned away (Baron, 1995). However, there have been problems with the device being able to work accurately in different home environments, and concerns that it might be costly to maintain (Sharot, 1994). To date, there is very little published information on its development.

Another advancement is a wristwatch with a microphone, a movement monitor and a thermometer that measure the wearer's physical reactions to indicate listening to or watching (Lindsay, 2000). Developed by the Swiss Company, Telecontrol, this technology apparently has been successfully tested in Europe. The cost of the watch is around NZ\$1500 and is currently worn by 22,000 Swiss each week.

The Portable Peoplemeter is yet another measurement device in its introductory stage in the United States. The PPM is a small device that uses in-built encoders to detect and store inaudible codes embedded by television broadcasters in the audio portion of their programmes. Panellists carry the meter with them for the day. At the end of each day, the participants place the meters into base stations that recharge the device and send the collected codes to a central processor. The data would then be processed to indicate the viewing of the participants. According to Purdye and Harvey (1994), this technology has been successfully demonstrated in Canada, although its introduction may require many more years of trialing. In the United States, the second and final trials of the Portable People Meter (PPM) were conducted in June, 2002 in the Philadelphia region (Market Research News, 2002). The potential for the concept behind the

technology is promising and, if successful, would eliminate many limitations of the push-button Peoplemeter system. As this technology is still in its infancy, it will be interesting to follow its future developments.

Clearly the latest passive Peoplemeter technology has the potential to register the presence of the viewer without requiring any effort from the participant. But it may still be years before its full use can be realised. The current form of Peoplemeter system will probably remain as the main measurement system for many more years.

2.2.2.3 Peoplemeter use in New Zealand

The present Peoplemeter system in New Zealand is operated by ACNielsen, who introduced the device in May 1990. The current system has a sample size of 470 households spread over 220 clusters with approximately 1150 panellists. The panel was assembled in two phases. The first phase was a nationwide stratified random sample of 1,760 homes where the head household shoppers were interviewed. At the end of the session, the respondents were shown brochures of the Peoplemeter system and it was explained that their household might be picked to take part (Danaher & Beed, 1993). At the second stage, one in four of these households was selected for panel recruitment. The panel was stratified by a three-way control matrix determined by four region levels, three household size levels (1, 2-3 & 4+) and three viewing levels (heavy, medium and light). At present, twenty-five percent of Peoplemeter homes are in Auckland, and ten percent in Christchurch. The rest of the panel is located around the country but their "whereabouts remain anonymous and concealed" (Lealand & Martin, 2001, p. 133)

Individual household members have their own number on the remote control handset. There is even a "pet" button for those who leave the set on to keep

their pet dog company while they are out (Zwaga, 1992a). Telephone checks on the audience compliance levels are conducted regularly. These checks are usually made in the evenings where panellists are asked what they are watching, and which channel they are tuned to. This information is then used to validate similar information about the household and their viewing choices from those obtained via the Peoplemeter.

The AGB McNair Peoplemeter in New Zealand is programmed to note any changes in set status that lasts for at least 15 seconds. A switch from one channel to another that lasts at least 15 seconds is recorded, while if the switching is all done within the 15 seconds, the change is not recorded.

2.2.2.4 Criticisms of Peoplemeter

The advent of the Peoplemeter system has not altogether overcome the criticisms of the rating data provided. There are three main criticisms of the Peoplemeter system. The first criticism questions the accuracy and representation of the data that forms the ratings for the programmes. The second criticism maintains ratings provide the "wrong" information about the quality of the viewing. It was suggested that ratings could not report on the "more critical factors of attention, interest and enjoyment" of programmes (Lealand & Martin, 2001). A third criticism is that the results provided by the Peoplemeter system focus on programme ratings rather than ad-break ratings. Yet this is often the "currency" used in the buying and selling of advertising airtime. These criticisms are now examined in more detail.

Worldwide Peoplemeter household panel sizes are small because the system is costly (Syfret, 1995). The cost of installing one in a household was between US\$3000 to US\$5000 in 1995 (McDonald, 2000). Only nine of the thirty-five Peoplemeter countries surveyed by Purdye and Harvey (1994) have sample

sizes of 1000 households or more. Because of the high capital cost of installation, the large expense of having to maintain the panel and the large data files they produce, sample sizes have to be kept relatively low. For this reason, all results are prone to sampling errors. This is particularly true for low-rated programmes and when the target groups aimed at by advertisers are small. For instance, in the Netherlands, where the sample panel consists of 2300, a rating of 10% would mean that 230 people were watching a programme (Boon, 1996). But if an advertiser is interested in a target group that is 30% of the sample, and the rating for the show is 5%, the size of the audience for this will only be 35 people ($0.05 * 0.3 * 2300 = 35$). Furthermore, as the number of channels has increased over the years, ratings per channel have dropped. This means that there will be even fewer viewers on which to base media decisions. The ratings for a programme that attracts, for example, 5% of the population might be based on as few as 50 households. In the United States, Ephron and Gray (2001) reported that the standard error of the Peoplemeter panel in the Philadelphia area ranged from 33% to 120%, largely due to its small panel size of only 270-panel households. These large sampling errors due to small sample size indicate a need for caution when interpreting the estimates in terms of precision.

There is also concern, over accuracy (in terms of bias) associated with the make-up and the maintenance of the Peoplemeter panel. Peoplemeter panels are difficult to recruit as most people find the device intrusive and the pushing of buttons a troublesome chore (Ephron & Gray, 2001). In the Netherlands, this non-response rate is as high as 70% (Boon, 1994). This could mean that the viewing of the panel would differ from the average Dutch viewing behaviour.

According to Lealand and Martin (2001), the Peoplemeter families recruited in New Zealand are "conservative in their habits and very fixed and predictable in their viewing choices". There may also be a poor representation of the diverse viewing habits of the multi-cultural television audience of New Zealand. However, even though sub-cultures may be represented, they would constitute a

small number; too small to be relied on to make good estimates. Lealand and Martin (2001) maintained, "the number and distribution of Peoplemeter devices in New Zealand cannot adequately represent the range or the diversity of the New Zealand audience" (p. 136).

Other limitations of the Peoplemeter are associated with maintenance of a reliable panel and control over its turnover rate. According to Gill (2000), panel members should be motivated to remain in the panel for several years, and incentives used to make sure they stay. If these factors are not well controlled, they can lead to bias. Milavsky (1992) reported that the turnover rate for the Peoplemeter panel in the US was estimated to be 62% annually. Of those who left, one-third left because of the agency's scheduled "deactivation" of panel households while the other two-thirds left on their own accord. Half of the unscheduled losses (30%) resulted from members shifting away, and another 30% simply decided to drop out. Milavsky (1992) warned that such high unscheduled turnover rates might bring about scepticism of the adequacy of the data, especially if the replacements who were recruited did not match the original intended sample.

Gane (1993) also commented that since the introduction of the Peoplemeter, there has been considerable variation in the implementation. For example, in the US, Nielsen (the research agency responsible for collecting the Peoplemeter data) does not use an establishment survey as the basis of sample selection; rather they use random probability selection to maintain the correct balance. There is also no standardisation to determine the panel area and how often the panel is rotated. In Europe, only natural rotation is employed. This refers to those who have moved, refused or are poor co-operators. In some countries (like the Chilean market), a 6-month rotation is used. As the television medium is used more internationally, there is an increasing need to be able to compare across different countries, which calls for greater harmonisation in the method used to measure television audiences (McDonald, 2000).

In the mid-1980s there was great scepticism about the ability of Peplemeters to provide audience data with sufficient accuracy. In its developmental stages, the Peplemeter system had difficulties measuring households that had video recorders or cable television. Billet (1985) commented that Peplemeters could not collect data fast enough to be able to track the changing patterns of video viewing. The current system has expanded and became more sophisticated in order to catch up with the complexity of electronic technologies, different methods of funding TV channels and the restructuring and deregulation of the media industry in many countries (Gane, 1993). Peplemeters can now identify whether the programme is being recorded and when the video is used in "play back". But it is still impossible to know what is being played back or whether it is a self-recorded or a pre-recorded programme (Buck, 1987). There are also problems related to homes with multi-set televisions, cable channels, households with TV games, and with out-of-home viewing (such as those in the pubs) that the system cannot measure.

Another limitation of the Peplemeter technology has been the issue of "compliance". This refers to the fact that it was often difficult to know for sure whether the viewer who claimed to be viewing the TV really was watching. This begs the question whether the panellists did indeed "push the buttons" on the remote control faithfully to ensure that accurate data were collected. Over time, there is also the worry that there may be a drop in button-pushing due to reduced interest. According to Gane (1994), the need for each respondent to push the button on their handset at the start of each viewing session and again at the end "... is still a source of panel burden and reporting inaccuracy which would be desirable to avoid" (p. 86).

Ever since the early days of the Peplemeter system, the problem of compliance has been dealt with by conducting "coincidental surveys" designed to test the validity of the data gathered. These studies involve interviewers contacting a

sample of households to determine whether the television was switched on and what viewing is taking place at the moment of contact. For each set, the respondent is asked who is watching and may be encouraged to validate the response by actually checking who is in front of the set. These results are then contrasted with the data collected via the Peoplemeter.

The method may be carried out over the telephone or during face-to-face interviews. There are two distinct ways of performing such a study (Kent, 1994). The first is to conduct it on the existing panel, while the second is to select a stand-alone sample not in the Peoplemeter panel. The two methods may have different objectives and provide differing insights into the validating of the panel measurement. However, despite the careful planning of such studies, different researchers have obtained differing (and sometimes contrasting) results from such surveys. Some of these studies are discussed in greater detail in the paragraphs that follow below.

Purdye and Harvey (1994) reviewed eight studies that addressed the issue of the compliance level of the panellists. The technique used for these coincidental studies was a telephone check that asked respondents to report their viewing at the moment of contact. This reported viewing was then compared with data collected by the Peoplemeter. In most of the studies, it was found that there was a range of between 73% to 92% compliance rate. Despite the high compliance rate, they suspected the methodology used overlooked the point that "most people (in the studies) are canny enough to deduce the purpose of the telephone calls..." and that "the bulk of the viewers whose buttons are not in "accurate positions" misreport their viewing..." to back up their earlier reports about their button pressing task. If this were true, the compliance levels in such coincidental studies would be over-estimated.

In New Zealand, Danaher and Beed (1993) were commissioned by the Television Broadcasters Group (TBG) "to study the level of compliance of the

panellists as part of the audit process required by the television networks in New Zealand". They conducted coincidental telephone survey during programmes over four time periods: 4.00-4.30 pm; 6.00-6.30 pm; 7.00-7.30 pm and 8.00-8.30 pm, with a stratified random sample of 400 panellists. They concluded that non-compliance did occur among their panellists but that it was "not serious". The compliance rate found in this study was 91.5% and supported similar studies in other countries. In addition, they commented that non-compliance did not seem to be related to the length of time on the panel. They suggested that overall programme ratings seem unaffected by non-compliance.

A similar study in Germany (Merz, Schumacher & Sudholt, 1993) also concluded that their Peoplemeter panel "provides a very accurate picture of TV usage". This study was commissioned to measure television ratings through an external sample by coincidental telephone interviews, and compared its results with those obtained by an internal coincidental check on a sample of the Peoplemeter panel. The study reported that there was considerable consistency between the external sample and the Peoplemeter panel.

But a common criticism of such validation studies is the way the interviews are conducted. In the New Zealand study (Danaher & Beed, 1993), interviewers identified themselves as representing AGB McNair, the company operating the Peoplemeter service. As a result there is a likelihood that the panellists felt they were being *checked on* by the agency that recruited them. They would tend to give responses that they thought matched the data recorded by the Peoplemeter device rather than their actual behaviour. Furthermore, the information provided by the respondent about other family members within the household might not be as accurate as about their own behaviour.

Another issue is that Peoplemeter data sometimes cannot prove that viewers who claimed they were viewing were actually watching their advertisements. Poltrack (1998) reported on the "phantom viewing phenomenon" in which he

suggested that a significant number of viewers of a particular programme (*Suddenly Susan*) were perhaps "not there" (both physically and mentally) despite it being reported as a high-ratings programme. This programme was screened in between two other highly rated programmes (*Seinfeld* and *ER*). From the ratings information provided, it appeared that 80% of the viewers of *Seinfeld* and *ER* were also watching *Suddenly Susan*.

However, results from a parallel custom tracking study that surveyed viewers via telephone interviews, found contradictory evidence. According to Poltrack (1998), this study showed that only about 32% stated they were "aware" of *Suddenly Susan*. An even lower percentage, just over 15%, stated they had actually viewed it. As the programme was placed between two highly rated programmes, Poltrack (1998) questioned whether this was a result of viewers "taking a break between *Seinfeld* and *ER* without logging out". Even if the viewers were there, the ratings data were unable to show the lack of "awareness", the viewers had of this programme. Moreover, only about 4.5% of the viewers indicated they "would watch" this programme in the future.

Interestingly, when the station moved *Suddenly Susan* to another time slot, the programme lost 60% of its viewers. Therefore, the ratings for this programme must have been "wrong" and even if the ratings were accurate, many viewers did not relate to the programme. According to Poltrack (1998), ratings data are quantitative in nature and cannot provide rich enough qualitative data that would give advertisers clear indications about the communicative effectiveness of programmes.

Anderson, Field, Collins, Lorch and Nathan (1985) also noted that children spent 13.4 hours a week in a room with the television turned on. This was lower than the national ratings provided by Peplemeters for viewers aged 2 to 5 years old, which averaged at 27.8 hours per week. As Peplemeter ratings record presence data, this study suggests that the viewing data for children were

over-rated. When attention (looking) measures were used, the results were different.

Another major criticism of the system is that it does not provide enough qualitative information about the viewers in front of the set (Baron, 1995). The system merely records "bums on seats" (Svennevig & Wynberg, 1986). It gives the characteristics and the number of the people in the room but not necessarily who in the household is actually paying attention to the set. Morley (1990) suggested that it is the varying levels of attention that would be of interest to advertisers and broadcasters alike. He sees Peoplemeter panels as producing data that do not reflect viewing. Perry (1995) also pointed out that the nature of Peoplemeter panels made it possible to measure only overall behaviour. When viewing differs in different circumstances, the panel was unlikely to reflect correctly all the viewing elements.

This concern over the lack of information is not new. For example, Lloyd and Clancy (1991) commented that information on the size of an audience reflected "... only exposure opportunities..." but failed to consider "...the effects of contemporary viewer behaviour...". The authors quoted a study by Television Audience Assessment, Inc. (1981), which indicated that 40% to 50% of a given programme's audience were engaged in all manner of distracting activities while "watching" and as much as half the audience left the room repeatedly during both programmes and commercials. In addition, more than 50% of viewers between the age of 18 and 30 watched more than one programme during an hour-long programme. Another 20% watched three or more programmes. Lloyd and Clancy (1991) concluded that the Peoplemeter systems could not "guarantee that viewers were actually "watching", partially distracted, fully distracted, or even in the room...". In other words, another major shortcoming of the Peoplemeter data is that they do not reflect the quality of the viewing (in terms of viewers' behaviour displayed when they were present during the programme).

For many years, starting from the beginning of the 1990s, Peoplemeter agencies have been testing "passive meters" in order to reduce "human errors" brought about by Peoplemeter panellists not correctly pushing their buttons. These passive devices were designed to detect automatically who was viewing as well as the channel being viewed without requiring respondents to push buttons or be consciously involved with any manual recording system. Moreover, Purdye and Harvey (1994) reported that although most countries are "moderately satisfied" with Peoplemeters, some of the countries (eleven countries out of thirty-four) studied would support a move to passive meter measurement. One reason given was that despite the efficacy of the push-button system, many countries had doubts that the present Peoplemeter system would be able to cope with the future problems caused by the increasing introduction of channels, a fragmented market (due to the audience now have more media choices, such as satellite and cable TV), and other "people problems" (related to the motivation of the panel to provide accurate data). However, the development of passive meters requires sophisticated software and other advanced technologies that would take time to develop and test. Even if they are successful, these improvements will only bring about a more accurate measurement of who is in the room not but will not identify more accurately the programme or advertisements being viewed.

Much of the previous work reported has focussed on providing ratings data for particular television programmes. However, the immediate problem for most media planners is not the ratings for programmes but the likely ratings for particular commercials screened in conjunction with particular programmes, for which programme ratings are only a surrogate. This is yet another major criticism of the Peoplemeter, because although, it may provide the ratings for the programmes, what is more important for advertisers is the number of people who saw their advertisements and whether the viewers were attentive or not to the advertisements.

It would be more beneficial for advertisers to know how many viewers actually reacted to their advertisements than be provided with information about the size of the audience. According to Boon (1996), direct responses about commercials will give more and better insight in the value of the money invested by advertisers than will ratings. Studies on direct responses include those that measure the effects of brand awareness and the change of attitude of viewers after viewing the advertisements screened.

Other types of studies directly probed viewers about what they were doing during the screening of commercials. In New Zealand, Colmar Brunton Research (1992), using a telephone survey, asked respondents about their viewing activities before the phone call. They found that "46.5% of the viewers reported physically leaving the room during commercial breaks, one-third talked or did something else, 11.7% zapped the set (changed channels or muted the sound) and only 8.1% watched and listened". These findings suggested the attention paid to commercials was generally quite low.

Such findings are however contested by Peoplemeter companies who have reported that the drop in ad ratings during programmes is often not very high, perhaps only about 10%, (Danaher, 1995). In a study carried out over November 1992, commissioned by the Television Broadcasting Group, Danaher and Beed analysed the data collected by 1000 panellists in the New Zealand Peoplemeter base and concluded that audience figures only drop by 4% on average during prime-time breaks. This figure was consistent with surveys conducted in the United States where there was a maximum drop of around 5% during ad breaks. Danaher and Beed (1992) also showed that viewers were exposed to 90% of the advertisements on the channel to which they were tuned before the ad break. In addition, they reported that for 91% of the time the TV set was tuned to the same station for the ad break, as it was 30 seconds before the break. This meant viewers switched channels less than 10% of the time.

In the United States, Cook (1995) gave details about a small-scale observation study conducted to collect data on push button accuracy and the eyes-on-TV screen data of their Peoplemeter panel. These eyes-on-TV data were recorded by passive image Peoplemeter technology that tracked the presence of the viewers and whether their faces were facing the screen. Observations included every person's presence in the room, their button pressing activities and eyes-on-TV screen status coded during a 30-minute "viewing study". The author reported that 95% of the audience in the room there for the programme were also there for the commercials. However, Cook (1995) also pointed that there was a great deal of other activity happening. Viewers were talking, reading, playing, doing chores and even asleep.

Despite the sophisticated technology it could be said that the Peoplemeter system currently used in many countries is still unable to record many qualitative aspects of ad viewing. Danaher (1995) concedes that some viewing behaviours cannot be measured by Peoplemeters. Such behaviour includes whether viewers are looking at the advertisements, muting the sound, or reading and talking during the viewing.

There has been much scepticism about the accuracy of the method by which ratings were obtained using Peoplemeters and if indeed Peoplemeters measure the "right" picture of television viewing. The system has many shortcomings, ranging from the small panel size used to represent a country's large and diverse population to the accuracy with which participants actually comply with the technology. The Peoplemeter measuring system may therefore be subjected to many human errors.

There is also the issue that the information provided is only a "slice" of the whole action in the television room. Ratings may report "who is present" and what they are doing with the television set, but television viewing is often accompanied by a variety of other activities, and viewers pay varying degrees

of attention to what is being screened. Being present does not therefore mean that attention to the material on television is guaranteed. Peoplemeter data cannot report on the attentiveness and different responses of the viewers to the screen (i.e. the quality of the viewing).

These points together suggest that advertisers are provided with information on the ratings of the programme (in terms of the size of the audience) but what is of greater importance is to know the qualitative aspects of how much attention was paid to the advertisements.

Most advertisers fear that lack of attention to programmes causes their advertisements to go unnoticed. The point is that measurement of the attention to the advertisements in which advertisers invest large sums of money to produce and air on TV, is often neglected in the information provided to advertisers. This information would be more useful for media planners than information on the ratings of individual programmes. This point is one of the main arguments of the present thesis.

2.3 Television Audience Viewing Research

The previous sections discussed the different audience measurement methods used in the television industry and examined their major flaws. One suggestion is that it may be unwise to rely solely on ratings, as they provide an incomplete and sometimes "wrong" measure of audience viewing. This is particularly frustrating when audience measurement data represent a vital currency central to the sale of advertising. There are, however, a number of studies that report on how viewers behaved during the ad breaks and whether there were specific factors that caused viewers to avoid seeing the ads. These studies are examined in greater detail in the sections that follow.

2.3.1 Introduction

Many social scientists have argued that it is important to study how people behave in their natural environment in their normal everyday lives. This view has been supported by researchers studying television-viewing behaviours (Ang, 1991; Lull, 1979). Such studies investigated the effects of television on the audience and how the viewers behaved.

Clearly, television watching is a complex behaviour and happens amidst a diverse range of conditions that either increase or distract attention from the screen. This should be of importance to television advertisers who are interested to know when and how the quality of their advertisement exposure was heightened or reduced.

There have been a number of ways of responding to this problem and these are reviewed in the next sections.

2.3.2 Peoplemeter studies

2.3.2.1 Introduction

Despite questions asked by media critics about the value of programme ratings (provided by Peoplemeter data) for advertisers, Peoplemeter agencies have repeatedly investigated ways to provide more descriptive data about viewers. They have realised the importance of providing information on the behavioural effects of the viewing for advertisers and of developing ways to measure responses to commercials. For example, Danaher (1995) investigated the factors that affected ad viewing, while Danaher and Lawrie (1998) reported how Peoplemeter data could be used as a surrogate behavioural measure of viewers' involvement in / appreciation of programmes.

2.3.2.2 Ads

Danaher (1995) analysed ad-break ratings for 325 ad breaks during prime-time programmes aired by three television channels in New Zealand. His overall conclusions were that ad avoidance by the New Zealand television viewing public is low and mainly done at random. In other words, viewers did not deliberately avoid advertisements. According to Danaher (1995), there were also few personal or household characteristics that correlate with ad-avoidance behaviours. In general, people of higher incomes, larger households and younger people tended to watch more ads than other demographic groups.

The audience attention was sustained particularly during the first ad of the ad break but thereafter, it dropped. At the mid-point of the break, some audience might return from switching channels or might have returned to the room. During the entire show, it was common to see similar variations in the ratings for the programme. Danaher (1995) suggests that such "comings and goings"

(such as viewers coming back to and going out of the programme or room) during the advertisement break were not unusual but rather part of the larger pattern of television viewing. He also suggested that viewers tended to switch channels during programmes and not during the ad breaks.

Furthermore, Danaher (1995) also reported that the proportion of viewers who were watching the programme and who also stayed on for the ads was high for most programmes. Primarily interested in the relationship between programme ratings and ad-break ratings, Danaher computed the ratings for ad breaks as a proportion of the programme in the programmes, a measure that he called ADRATIO. The formula for this calculation was the ratio of ad-break ratings to programme rating.

$$\text{ADRATIO} = \frac{\text{Ad break rating}}{\text{Programme rating}}$$

Ad breaking rating = average rating over all the ratings in the ad break

Programme rating = average rating for programme minus NPM

NPM = station promotions, public service announcements and commercials (Non-programme material)

The results of that study showed the average ADRATIO to be 95%. This suggested that during the programme, the proportion of viewers who had "dropped off" (not viewing) during the ad break averaged only about 5%. In terms of the relationship between this ADRATIO and programming, the study also reported that programme duration and programme type may affect the way viewers react to advertisements. The results of ADRATIO by programme type is reproduced in Table 1.

Table 1: ADRATIO by programme type

Programme types	ADRATIO
Game / quiz	101.4
Soaps	95.5
News / documentary	95.1
Comedy	94.5
Sports	94.4
Drama	94.3
Movies	89.6

Shorter programmes such as game shows, which were usually half an hour long, had higher ADRATIOS than movies, which were normally aired for more than two hours in length. In other words, ADRATIO varied by type and length of programmes. But Danaher argued that it would be difficult to separate the effect of programme duration and programme type since most programmes usually have "fixed formats" and time constraints. However, he reported that Movies had the lowest ADRATIO and Soaps the highest. Soaps tended to have high "audience loyalty" within the show as well as across episodes. Movies, on the other hand, tended to attract "grazers" from other channels who were channel-surfing (switching) during the ad breaks and had stayed on to watch the programme.

In addition to ADRATIO, Danaher (1995) also investigated the actual number of minutes viewed during the ad break. The author developed a measure termed as PROVIEW, which was calculated as:

$$\frac{\text{Number of seconds viewed during the ad break}}{\text{Number of seconds of potential viewing during the break}}$$

The author found that, on average, PROVIEW for households was as high as 90.5%. He concluded that ad avoidance was a random activity not common during ad breaks. Ad avoidance increased mainly because people recorded

programmes on the VCR and tended to "fast forward" the advertisements during playback. Households with higher incomes also tended to avoid the advertisements more often than other income groups. In addition, viewers also tended to avoid ad breaks that were long and consisted of many ads. The author concluded that while viewers did avoid ads and there was a drop in the ratings during the ad breaks compared with the programme ratings, this drop was not alarmingly serious (shown by ADRATIO for programme types).

Danaher's (1995) overall conclusions were that ad avoidance by the New Zealand television viewing public is low and mainly done at random. In other words, viewers did not deliberately avoid advertisements. There were also few personal or household characteristics that correlate with ad avoidance behaviours. The only ones noted were that younger people and people of higher incomes avoided ads more often than other population groups. Larger households watched fewer ads while non-European New Zealanders and rural dwellers tended to watch more ads.

Although this study shed some light on the relationship between ad viewing and programmes, the data was still based on the size of the audience and not on actual details of viewer behaviour activities. Even though there was information on what ads were effective, it was based on the number of viewers who had pressed a button on the Peoplemeter device and not based on whether they were actually looking at the ads. Even when viewers returned to the programme or the room, it was only an assumption that they were watching. In reality, they might be in the room but doing something other than paying attention to the ads. The instances of channel switching may under-report ad avoidance as these findings may only represent a small proportion of viewers who chose to change channels. There are other ways to avoid ads. There may be a larger group of viewers who did not switch channels but were engaged in other competing activities that took their attention from the TV. In conclusion, it would be better to know the actual types of viewing behaviours that happened during the ad

breaks rather than know how many people had been present during which parts of the ad break.

2.3.2.3 Programmes

In Danaher and Lawrie (1998), the main objective was to use Peoplemeter data to develop behavioural measures of programme involvement and appreciation. For that study, the authors developed another two measures: PMV and P80+.

PMV represented the percentage of minutes viewed and was calculated by:

$$\frac{\text{Number of minutes programme viewed}}{\text{Number of minutes programme aired}}$$

The PMV score for the whole sample would be the average proportion of minutes viewed by all viewers. However, a more stringent criterion was that a viewer must have at least viewed 20% of the available minutes to be included in the average. Another measure was therefore developed (P80+). The measure P80+ represented the proportion of viewers who watched 80% or more of the programme. The authors argued that this would show the viewers' commitment to the show.

Danaher and Lawrie (1998) reported that there was high appreciation for most programmes: PMV averaged 78% and P80+ averaged at 60%. These commitment scores were also consistent across episodes of the same programme types, except for comedies and magazine shows. There was also positive correlation between programme ratings and commitment scores. Popular shows had more viewers and the viewers (measured by PMV and P80+) were more committed. The correlations were also higher for information type programmes than entertainment ones. In particular news and soap operas had higher than average commitment scores than other programme types.

Although Danaher and Lawrie (1998) stressed the importance of using behavioural measures, their study relied on presence data (i.e. the number of seconds or minutes viewed). As the study did not specifically examine ad-viewing behaviours, there was no guarantee that high PMV and P80+ scores meant high ad attention.

2.3.2.4 Summary

According to the proponents of the Peoplemeter system, research on ad-viewing behaviours examined by Peoplemeter data have shown consistent results that ad avoidance was not a serious matter for advertisers. In most cases, the ad ratings were just as high as the programme ratings. Ad avoidance was mainly random and there are a few personal, or household characteristics or programmes that correlate with ad-avoidance behaviour. These factors may be age, income levels and location of dwelling. There was no particular reason why people chose to watch ads or not but it was common that viewers switched from one channel to another or left the room during the ad breaks. Ads placed at the beginning of the ad break were more often "watched" than the other parts. In addition, alternative TV viewing devices such as the VCR and remote controls meant there were increased opportunities for viewers to avoid ads. Nonetheless, channel switching was mainly carried out between programmes rather than during the ad breaks.

With regard to programme types, viewers of Soaps had the highest proportion of ad viewers and high audience loyalty. But programme type must be considered in relation to programme duration. It was difficult to separate the effects of programme types and programme duration. Nonetheless, most high ratings programmes have high ad ratings. In fact the drop in ratings during the ad break from the programme rating in most cases was very small (5%).

Peoplemeter research also showed that popular shows had more viewers and that these viewers were more committed to the show. Among programme types, news and soaps have higher than average commitment scores. In terms of age differences, older viewers showed higher programme commitment than younger viewers.

But Peoplemeter research on viewing behaviours suffers from the limitations imposed by its technology and methodology. Some of the criticisms of Peoplemeter technology are also relevant here; there are doubts that the data accurately reflect ad-viewing behaviour. Although pertinent and informative details about the ads and viewers' involvement levels for different programmes were given, the data were derived from estimates of the number of people present and assumed to be watching. As the information was based on presence data, the question is also whether presence could be a substitute measure of attention (paid) to the ads and how truly it reflected what the viewers did during the ad break.

2.3.3 Observational research

Television viewing is a complex process that includes multiple activities and varying levels of attention. Viewers could range from attentively viewing to engaging in competitive behaviours (such as reading, talking or sleeping) during which they are not at all orientated to the set. Alternatively, they could also be engaged in complementary activities (such as eating), where they could be simultaneously engaged in more than one activity during the viewing. Such data on different viewing patterns often go undetected by Peoplemeter research but could easily be measured by observational research.

While Peoplemeter research shows data about which channel was being watched and how many viewers were present, a number of other studies report

the activities of the viewers during the ad breaks. For example, there are studies that investigate the amount of attention viewers paid to the TV screen, and how programme involvement affected viewing and viewers' responses to ads. These studies can be classified into three categories: studies where observers recorded viewers' behaviours watching TV; in-home filming where the researchers were not present; and studies where viewers were asked to rate their perceptions of programmes or their behaviour on attitude scales. The observational studies are examined in the following sub-sections.

2.3.3.1 Historical overview

This section provides a historical review of the major research that used recording devices to collect data on television audiences. The methodology used in these observational studies is described here and their key findings are discussed in the next section (see 2.3.3.2).

The earliest of the recording methods used in audience research involved setting a camera near a television and periodically taking pictures of the audience. Allen (1965) reported on four "Dynascope" studies done in Oklahoma and Kansas that used photographs to observe television viewing at home. This was conducted in a total of 95 homes for two weeks. These studies covered the period from 1961 to 1963. The Dynascope took pictures of the normal family viewing area on 8 mm, 16 mm or 35 mm film. Another study, carried out by Bechtel and Akins in 1972 (reported in Collet, 1986), used two cine cameras placed in the room. One faced the viewing room and the other faced the television. They studied 20 families in Kansas City, Missouri, in 1970 over a 6-day period. The videotapes, recorded in real time, were coded according to the amount of time people attended to the television. In 1985, using videotape, a pioneering study was carried out by Anderson, Field, Collins, Lorch and Nathan in the United States. They measured children's attention to television by using

time-lapse video camera technology to record their viewing behaviours. These studies show low levels of attention paid to television.

Another important development of the technology for audience recording is the so-called C-BOX developed by Collett at Oxford University in 1986. The C-BOX was a cabinet that housed a television, a half-concealed camera situated above the TV, and a video-recorder. The video recorder also recorded the television picture and soundtrack. In this way, it was possible to see what was happening in the room as well as what was screening at that time. This information was provided as a miniaturised inset on the tape. The whole equipment was tested in 20 homes in the London area. Sample households also completed diaries while the device recorded their actual behaviour. Coders then assessed the amount of time each adult in the household had been present in the room during each programme viewed, and the amount of time they had spent looking at the screen.

The C-Box provided long-term continuous observation that would measure both presence and attention as well as provide information about the type of programme in which commercials are being watched, the characteristics of the audience who watch commercials and other population parameters. The behaviour of the viewers could be passively taped without requiring the audience to comply with any equipment measures. Since there was no researcher present during the taping, the effects of interviewer bias were also reduced. Since the respondents were not required to take any action such as pressing buttons over time, the problem of respondent fatigue was also avoided.

Then, in 1987, Svennevig, funded by the IBA (Independent Broadcasting Association, UK) reported the use of H.O.M.E. (Home Observation and Monitoring Equipment). The equipment was a further extension of Collett's "C-box". This device used two VCRs for the recording: one for the camera and one for the television picture. Only nine North London households participated in

this study. This study provided descriptive transcripts of the tapes and broad generalisations about viewers and their use of television.

Gunter, Furnham and Lineton (1995) produced further more detailed findings of the IBA observational research conducted by Svennevig (1987). According to Gunter et al. (1995), the original analysis was superficial and impressionistic and more systematic analysis of the data was necessary. In total, 230 hours of videotape were coded and analysed. The researchers analysed the proportion of time viewers were present in the room while the TV was on and the amount of attention given to the screen while it was on. They also reported on the distribution of time present and spent looking at the screen across different television channels and different programme types, and provided an account of other activities performed during the viewing.

2.3.3.2 Viewing patterns

Using the methodologies in the previous section, one common conclusion from observational studies is that even when the set was switched on, it did not mean viewing was taking place.

Allen (1965) reported that on average, depending on the type of programme, the television set was ignored between 25% and 50% of the time it was switched on. Besides supporting the findings of Allen (1965), Bechtel, Achelphol and Akins (1972) also confirmed the notion that there was often "no audience" in front of the television. According to Gunter, Furham and Lineton (1995), about 50% of the time the television was on there was no one present, and even if they were, they did not look at the set all the time. Anderson, Field, Collins, Lorch and Nathan (1985) also reported that the amount of time children were observed to be "looking" at the set was around 67% of the time the television set was on.

Collett (1986) commented, "presence cannot be equated with attention". According to this study, on average, the amount of time someone was in the room when the television was on was about 80% of the time. But the amount of time people spent "looking at the screen" was only around 65%. This meant a proportion of viewers were "present" but not paying attention to the screen.

Moreover, according to Svennevig (1987), viewers often "left and returned" to the programme. This study reported that, although, on average across all his families, viewers were present for around 70% of the time, viewers paid varying degrees of attention to the TV screen. They divided "attention" into two categories, "full" or "partial". Full attention was when viewers were judged to have paid total attention to the TV. When there were other forms of involvement besides full attention, it was judged to be "partial attention". There was also "full inattention", when viewers were judged not to be attending at all. According to the authors, full inattention accounted for 33% of the viewers time, while viewers were partially attending around 20% of the time and, viewers paid full attention around 27% of the time.

Television viewing is also often accompanied by activities other than "pure watching". Allen (1965) found that 25% of the time when the set was turned on, viewers were engaged in activities other than watching the TV. Anderson, Field, Collins, Lorch and Nathan (1985) reported that children spent one-third of the time doing things in front of the television other than looking at the screen.

Zwaga (1992a), using the video recording method, also used a C-Box to study eight households and obtained 275 hours of video taped television viewing. His custom-made cabinet contained a 21" television, a low-light camcorder and several video recorders. Although he was interested in the relationship of television and the social life of viewers, he claimed that television was a form of "moving wallpaper" with viewers tuning out as soon as the commercial

breaks come on. He analysed the tapes qualitatively by merely monitoring the families' actions, and produced descriptive viewing "vignettes" about each family. He reported case study extracts of the family-viewing day.

Zwaga (1992b) commented that television watching could not be considered as a "prolonged, sustained effort requiring full attention" (p. 13). During ad breaks, TV viewing "oscillates between it being a primary and secondary activity". He concluded that while programmes were watched more attentively, ad breaks were signals for talking and doing other things. Families often avoided ads by switching channels or by totally ignoring them. Viewers tended to tune out as soon as the ad breaks came on. Clearly there are many different opportunities (for a change of action) open to the viewer as soon as the advertisement break comes on.

Examining the effect of programme type on viewing behaviour, according to Betchel, Achelphol and Akins (1972), viewers paid more attention to movies and factual programmes than other programmes. Betchel, et al. (1972) reported that during the time viewers were present, for movies, there was at least one person looking at this programme 76% of the time. Viewer attention to other programme types was: children's programmes (71%), suspense series (68%), religious programmes (67%), family series (66%), game shows (66%), talk shows (64%), melodramas (59%), sports events (59%), news (55%), and commercials (55%).

However, Svennevig (1987) reported that soaps and serials attracted the largest audiences but viewers paid little attention to these programmes even though they were present in the room. By contrast, viewers spent less actual time watching factual programmes, but while they were present, they watched more attentively. This was also supported by Gunter, Furham and Lineton (1995), who reported that factual programmes such as news, documentaries and religious programmes "commanded" the greatest visual attention. However, the

viewing behaviour tended not to be continuous and viewers were looking away from the screen for substantial periods of time the programmes were shown. These findings again highlight the point that being present in the room did not mean that there was continuous attention. The question is at which point during the programme did the attention decline. Furthermore the relationship between "being present" and "looking" may also vary across different programme types.

Visual attention was also found not to correlate with the total amount of time viewers spent watching television. This means that both heavy and light viewers of television could be just as attentive to programmes and commercials. According to Anderson, Field, Collins, Lorch and Nathan (1985), heavy viewers (those who spent large amounts of time with the television) did not necessarily spend more time paying attention to the screen. For instance, their heaviest child viewer spent 39.8 hours per week with the TV turned on, and yet was only visually attending 3.4 hours per week to the screen.

Attention has also been shown to vary with age and gender. In comparing children and adults, Anderson, Lorch, Field, Collins and Nathan (1986) reported that children spent about an average of 12.8 hours per week in the room with the TV turned on and adults spent around 11.5 hours. Adult females paid significantly less visual attention than children, 54% versus 70%. Male adults on the other hand, spent around 63% of their time watching television. In terms of visual attention paid to the screen (measured by the frequency of times they had looked at the screen), the authors found there was a large increase in percent visual attention from ages 0-5 years (averaging around 70%), which levelled out during school years and then decreased during adulthood.

Betchel et al. (1972) had also reported such a declining trend in attention among adults but further commented that housewives were most inattentive to the screen. Anderson et al. (1986) suggested that for adult viewing, television was an easily learned activity and could be time-shared with other on-going

activities. Less visual attention was needed for full comprehension, especially since the correlation between looking and listening to television decreased with age. These findings suggest that different viewers (of different ages and gender) used different strategies when watching television, and different viewers experienced different forms of involvement with the media. Anderson et al. (1986) concluded that attention to television varied significantly with the content of the programmes and with age and individual viewing styles.

2.3.3.3 Relationship between programme and ad viewing

Krugman (1983) suggested there is a general consensus amongst researchers and industry practitioners that television viewers (generally) pay more attention to programmes than to commercials. But Krugman, Cameron and White (1995) also pointed out that many of the studies on visual attention tended to report more on programme-viewing activities rather than on commercials. They suggest that much more information about on-the-spot viewing activities during the commercial break is needed to substantiate research on the quality of viewing and advertising responses.

One early study, Steiner (1966), employed in-home observers ("spies") to examine the amount of time viewers spent watching television. He reported that while 76% of the viewers paid full attention to the programme just before the ad break, only 47% of the viewers paid full attention to the ads. Another study, Bechtel, Achelpohl and Akins (1972), reported that while 76% of the viewers watched movies, only about 55% watched the advertisements in the programmes. More recently, Abernethy (1991) concluded that viewers were "exposed" to ads only 68% of the time while ads were screened and by comparison with programmes, this exposure was much lower. These studies clearly show there is a difference between attention to programmes and attention to ads.

In a study using in-home participant observers, Moriarty and Everett (1994), reported that ad breaks stimulated viewers to change channels (91%) more often than programme content (50%). In addition, they reported that viewers avoided advertisements more frequently by "physical avoidance" such as ignoring the ads, leaving the room or talking (58%) than by using "mechanical avoidance" such as the remote control to switch channels or mute the sound (53%). By contrast, viewers who were uninterested in a programme were more likely to use the remote control than physical avoidance. The authors suggest that ad viewers tended to create a greater psychological distance between themselves and commercials than between themselves and programmes. These findings suggest the viewing behaviours differed during ad breaks and during programme viewing. The next question is whether programmes encourage viewers to watch ads or not.

A study conducted by Krugman and Shamp (1992) using in home observers examined the relationship between programmes and ad viewing. They measured "visual orientation" as the proportion of eyes-on-screen (EOS) time viewers spent with TV. Of subjects who were visually oriented to the programme, 62% were also visually oriented to the set during the commercial break, whereas viewers who were only partially or not visually oriented to the programme paid less attention to the advertisements. These observations suggest that there is a strong correlation between programme and ad viewing.

In another study, also employing in-home observers recording the behaviours of viewers as they watch television, Krugman, Cameron and White (1995) found an average of 33% eyes-on-screen time for commercials as compared with 62% for programming. Less than 50% of the viewers had their eyes on the screen more than 50% of the time. The authors explained that the drop in the attention paid to commercials was mainly due to today's multi-channel television environment and the increased use of the VCR and the remote control. The suggestion is that viewers tended to switch channels during the ad breaks.

In addition, this study also concluded that viewers engaging in more "competitive" activities, such as reading, hobbycraft and conversations not related to the programme, during the programme paid lower attention to both programmes and advertisements than other viewers who were engaged in less competitive activities. This further supports the suggestion that when attention is more concentrated during the programmes, attention to the ads may be higher than when the concentration is lower. This reiterates the point that attention to programmes positively influences attention to the ads.

In summary, the above observational studies have indicated that attention to advertisements may be lower than attention to programmes. These studies also reported a positive relationship between programme attention and ad attention. Viewers who paid high attention to the programmes are more likely also to pay high attention to the ads.

However, the question is whether ad viewing is a simple function of programme viewing or whether there are more complex factors involved. According to Bogart (1989), the surrounding programme content is rarely a neutral backdrop (p. 89). Thus viewing of commercials must be considered in the context of the programme that delivered them. The suggestion is that some programmes are "better vehicles" for advertising than others.

Over the years, different studies have emerged that have tried to determine the factors contributing to favourable programme contexts that encourage favourable advertising viewing. Many of these studies investigated how programme types which differed in content, impact and attractiveness influenced viewers' recall and responses to the advertisements. However, researchers have not been able to identify what makes one programme context "good" or "bad" (Norris & Colman, 1994).

It has been suggested that some programmes create an overall impact on viewers and that this impact often affects the way viewers evaluate advertising. For example, Broach et al. (1995) reported that evaluations of advertisements were enhanced after pleasant programmes and depressed after unpleasant programmes. Goldberg and Gorn (1987) suggested that commercials viewed in a happy programme were perceived more positively, were recalled better, and even elicited a higher purchase intent than those seen in sad programmes. Furthermore, Kamins, Marks and Skinner (1991) found that those commercials classified as happy placed in programmes that were classified as happy, were more well-liked than the same ones placed in programmes classified as sad. To make the distinction between a happy and sad ad, the researchers subjectively pre-selected two set of commercials they thought represented "upbeat and happy" commercials and "sad and depressing" ones. These commercials were then tested again by respondents who rated the commercials on a five-point likert scale on the dimensions of happiness and sadness. Their conclusions were that a sad commercial was better liked in a sad programme than the same commercial viewed after exposure to a happy programme.

Research done by Schumann et al (1990), Murray, Lastovicka and Singh (1992) found that placing advertisements in liked programmes resulted in more positive product impressions. Viewers formed more positive thoughts about the product when they watched programmes that made them feel good. They suggest that programme liking was a high involvement condition that influenced viewers to respond positively to the ads.

These findings show that viewer involvement with programmes influenced their assessment of the advertisements in the programmes. It may be that a positive mood evoked during the programme brought about a corresponding positive feeling for the ads. There is a possibility that feelings (such as liking) elicited during the programmes could be "transferred" to the advertisements, especially when the viewer involvement with the programme is high. Thus, the factors

that influence programme involvement may have a corresponding effect on the response to the ads.

2.3.3.4 Summary

Various ways have been used to study viewing patterns of ad viewers. Observational research methods (Allen, 1965; Betchel et al. 1972; Collett 1986; Andersen et al. 1986; Svennevig 1987; Krugman et al 1995; Gunter et al. 1995) have proven to be more appropriate in collecting such data than Peplemeter studies. An important issue is that being present in the room does not tell enough about the actual behaviour of viewers and the attention paid to advertisements. Attention has also been known to vary across different programme types, as a result of time of day, sex and age. It is known that presence does not necessarily correlate with attention. These two concepts (presence and attention) need to be measured simultaneously.

Attention to ads has also been found to correlate with attention to programmes. Factors that involve viewers during programmes may also affect viewers watching the advertisements in these programmes. But there are opposing views about how programme involvement affects ad involvement. Some researchers maintained that high involving programmes are more effective for advertising than low involving ones. Others maintained that ad viewers who are moderately involved react more favourably to advertising. Differing results may be due to differences in the methods used to measure involvement.

2.3.4 Effects of programme impact on viewing behaviour

2.3.4.1 Introduction

Over the years, programmers and television producers have added to the complication of programming by producing many different types of programmes intended to keep their audiences entertained and to satisfy a whole range of viewers' tastes and needs. In the field of media and advertising research, programmes can be grouped by their general content and characteristics (e.g., situation comedies, sports, etc.), or by viewer demographics (e.g., children, teens, adult programmes, etc.).

A refinement of genre classification is that developed by Barwise and Ehrenberg (1988) where programmes are differentiated in terms of the demands they make on the viewer. Barwise and Ehrenberg's (1988) main contention is that rather than there being many genres, these reduced to two categories namely categories *demanding* and *less-demanding* (also referred to entertainment programmes).

An additional approach uses (e.g., Hoffman & Batra, 1991; Clancy and Lloyd, 1999) personal evaluation of programmes to determine the level of involvement a programme generates. Often this involves a self-completed questionnaire where viewers give their responses immediately following the viewing of the programme. This last method differs from those based on genres because it is based on a personal, subjective evaluation of specific programmes. It is entirely based on viewers' responses, and researchers (Hoffman & Batra, 1991) do not make any assumptions about the programmes or their content.

2.3.4.2 Demandingness according to Barwise and Ehrenberg (1988)

Although not explicitly studying television advertising, Barwise and Ehrenberg (1988) found differences in the viewing patterns between what they classified as *demanding* and *less-demanding* (entertainment) programmes. *Demanding* programmes were defined as ones that involve the viewer and required the viewer to put in effort to watch them. Barwise and Ehrenberg (1988) developed these two major types of programmes based on the effect programmes have on the viewers. They measured the demandingness of the programmes by getting viewers to indicate if each programme either "helped (them) think" or "helped (them) relax". Based on this, informational programmes (news, public affairs, documentaries), heavy drama and cultural programmes were termed *demanding* programmes. Entertainment programmes such as light entertainment, light drama, films and sports programmes were *less-demanding* programmes. According to Barwise and Ehrenberg (1988), about 10% of the viewers said entertainment programmes "make them think" and 55% said that they "helped them relax". For *demanding* programmes, about 60% of viewers said they "make them think", and 15% said they "helped them relax".

Barwise and Ehrenberg (1988) were interested in the way TV was considered in the UK. One of their major findings was that although the tastes for television programmes differed widely, viewers generally preferred to watch a variety of programmes of diverse types. Using programme ratings (presence in the room with the set on) data as the basis for their research, they reported there was little variation in the way different groups of people divided their viewing across different programme categories. However, viewers in Britain spent 60% of their time on entertainment programmes and around 30% on *demanding* programmes. Barwise and Ehrenberg (1988) suggest the proportion of viewing for the two categories did not vary with income levels, sex differences and age. Moreover, this proportion also did not vary by user type. Light or heavy TV viewers spent the same proportion of their viewing time on *demanding* and

less-demanding (entertainment) programmes as the rest of the sub-groups. The implication is that this assessment of *demanding* or *less-demanding* programmes did not vary with viewer type.

As a result, Barwise and Ehrenberg (1988) suggest the television market could not be segmented. According to them television has a universal or mass appeal, and they concluded that TV advertising aimed at specific target groups would be unlikely to be very cost-effective. However, despite television audiences being unsegmented, they may still react to different programmes differently. Viewers watched a variety of programmes; some demanding and many relaxing ones. Programmes that demanded deep instructive and emotional involvement were described by Barwise and Ehrenberg (1988) as "high impact". Generally, people tended to like watching "high impact" demanding programmes more than entertainment programmes and this did not depend on how often viewers watched them.

The authors claimed that earlier research (Barwise, Ehrenberg & Goodhardt, 1979; Ehrenberg & Goodhardt, 1981; Meeneer, 1987; Stewart, 1981) also reported similar findings. Although fewer viewers watched *demanding* programmes, these programmes had a longer lasting effect on the viewers. Viewers of these programmes expected them to be more interesting and/or enjoyable than the more relaxing ones. Otherwise, the viewers would choose to watch a *less-demanding* (entertainment) show or not watch at all.

The authors also suggest that a *demanding* programme had more frequent viewers who liked it better than some other programmes. Normally, programmes that were watched less frequently (measured by the number of episodes watched in the series) were less well-liked by their viewers. In particular, low-rating entertainment (*less-demanding*) programmes tended to get lower liking scores than programmes that had more viewers. But within the range of *demanding* programmes, there was no clear (Double Jeopardy) effect.

Although there was a wide variation between individual programmes, high-rated *demanding* programmes tended to be as well liked by their respective viewers as lower-rated *demanding* programmes were. In other words, in different situations, *demanding* programmes still have a higher liking impact on viewers than the *less-demanding* ones.

The explanation was that only viewers who were genuinely interested in these programmes watched them. Watching these *demanding* programmes also required more intellectual and emotional effort, and viewers involvement was heightened. This in turn made the viewing experience rewarding for the viewers. The suggestion is that demanding programmes differed from the more relaxing programmes because of their ability to impact on viewers, and that viewers sought different levels of gratification from the viewing experience.

Barwise and Ehrenberg (1988) argued that the requirements of advertisers, broadcasters and viewers differed. While viewers wanted variety (of programmes) in their viewing, advertisers placed their ads over a few programmes but wanted the largest proportion of viewers to see them. Viewers, however, did not restrict their viewing only to popular shows. In fact according to the authors, viewers only watched two or three of the top ten programmes in the week. As a result, many less popular shows had been pulled from the weekly schedule at the expense of viewer choice. As *demanding* programmes tended to have smaller audiences, broadcasters often replaced them with higher ratings (entertainment) programmes. This example illustrates that there are limitations in the use of ratings data. Clearly, if *demanding* programmes have more attentive viewers who pay more attention to the ads shown during them, this would not be detected by ratings data. Perhaps other measures of viewing (such as viewing behaviours) may be warranted in order to investigate more fully the impact programmes have on ad viewers. While ratings data suggest that it was not viable to segment the television audiences, it may be possible to do so on evidence of how viewers behave during certain programme types.

In conclusion, Barwise and Ehrenberg (1988) distinguished between programmes on the basis of how much viewers "invested" in the programme (in terms of mental and emotional effort). *Demanding* programmes tended to have greater impacts on viewers but it is not clear how they would affect actual viewing behaviours. The authors, of course, were not examining attention (and their data were based on presence). Instead, they were describing the effects of "liking" and "repeat viewing" of programmes in two categories of viewing situations: *demanding* versus *less-demanding*. Nonetheless, how viewers react to these programmes would have important implications for advertisers.

2.3.4.3 Involvement

A number of studies suggest that the level of involvement a viewer has with a programme will affect the way ad messages are received (Bryant & Comisky, 1978; Lloyd & Clancy, 1991; Park & McLung, 1986; Soldow & Principe, 1981; Thorson & Reeves, 1986; Hoffman & Batra, 1991). This observation has prompted some researchers to focus on quantifying viewers' involvement with individual programmes rather than relying on genre derived classifications.

Hoffman and Batra (1991) suggested that Barwise and Ehrenberg's (1988) categories serve as a good starting point but that more empirical work was needed. They pointed out that most media research tends to equate involvement with viewer arousal, and suggested there was a need to study programme involvement as a multi-dimensional construct that should include cognitive and affective factors. They criticized earlier studies as manipulating the viewers' response by showing viewers programmes that were already pre-classified by the researchers. They proposed "programme impact is the measure of how much the viewer must give to the programme in order to get something back" and argued that it was therefore important to measure the viewers' actual

response to the programmes. They developed a rating scale that measured how viewers were involved with programmes cognitively and affectively.

Their analysis of viewers' personal evaluation of 39 programmes led them to believe that programmes involve viewers both in how they felt as well as how they processed information to suit their own purposes. In a study using several different programme types, the authors obtained individual ratings via phone interviews of programmes viewers had just watched. Their questions included asking viewers how they felt individual programmes had impacted on them, how each programme rated among other programmes, and how viewers graded these programmes. Using the viewers' evaluation of the programmes and canonical discriminate analysis to examine the dimensionality of programme impact, the researchers then grouped the programmes into three categories: High Cognitive, High Affective and Low Impact programmes.

High Cognitive programmes were programmes that viewers "watch to learn from" and "to get more from them". High Affective programmes were those that "touch (viewers') feelings". They noted, however, that some High Cognitive programmes also touched the feelings of viewers. In contrast, Low Impact programmes were those viewers "watch to pass time", from which they "do not learn anything" and which "do not touch (their) feelings".

However, they did not fully elaborate how the variables of cognitive and affective impact would "involve" the viewer simultaneously. It was unclear how specific programmes were classified; some programmes were described as belonging to more than one category (e.g., a particular programme could be high affective as well as high cognitive). Moreover, it was unclear how the authors came up with two categories of "high impact" programmes (i.e. High Affective, High Cognitive), and just one label for all "Low Impact" programmes.

According to the authors, High Impact programmes are those that affected viewers either affectively and cognitively or both affectively and cognitively. In general, High Impact programmes were described as enabling the viewers to get more and learn from them, while Low Impact programmes mainly aroused their feelings.

In addition, viewers of each programme type were described as displaying varying degrees of attention to the programmes and were involved with the programme somewhat differently. Three main measures were used in their evaluation of the viewers' behaviour: "attention"; "if viewers left the room (presence)"; and "talking". Four levels of attention were also included in their analysis: "all attention"; "most attention"; "some attention" and "low attention". If viewers left the room, this was measured by when it had happened: "only during the advertisement"; "only during the programme"; "at anytime whether during the advertisement or the programme"; "that they did not leave at all". A third category, "talking", included whether the talking was "about the programme" or "about something else".

Hoffman and Batra (1991) concluded that viewers gave the highest level of attention to High Impact Cognitive programmes, and were less likely to talk or leave the room during the programme. Viewers watching High Affective programmes were reported to have paid some attention to the screen and left the room during the programme and the advertisements. When they talked, it was about the programme. The authors suggested viewers left the room during the advertisement break to free themselves from the emotional tensions built-up during the viewing of these intense programmes. In contrast, viewers of Low Impact programmes paid little attention to the screen, talked about other things besides from the programme, and were more likely to leave the room during the programme itself.

The authors suggested the probability of the viewer staying in the room and watching the advertisement would be highest during High Cognitive programmes and lowest during High Affective programmes. The probability of viewers seeing the advertisements during High Affective programmes was not very high as viewers tended to leave the room during the ad break to release the emotions aroused by the programme. As for Low Impact programmes, viewing behaviour was "random". The authors suggested these viewers were unlikely to assimilate the programme or the advertisements.

Lloyd and Clancy (1991) also examined viewers' personal evaluation of programmes to measure the programme involvement level of viewers. They showed participants four different programmes that carried the same advertisements. They were all popular prime-time weekly shows and included a comedy, a drama, a mystery and a Sunday evening news. After the viewing sessions, viewers rated the programmes on a self-rating scale based on 30 items. A factor-analysis of the data uncovered three dimensions which they labelled: programme involvement; values incongruence; and mood enhancement. The first of these dimensions, programme involvement, measured the ability of the programme to influence the emotional and informational impact on viewers and the viewers' levels of interest, involvement and concentration. This dimension encompassed the "entertainment" value of programmes that viewers sought and that involved the viewers. It measured how viewers found "personal connections" with the "stimulus" in the programmes such as empathising with the main characters or "being touched" by the events in the show. According to the authors, this "entertaining" value of the programmes had the highest degree of reliability and validity among the three dimensions in measuring viewers' involvement.

After obtaining the level of involvement of viewers with each type of programme, Lloyd and Clancy (1991) then analysed viewers' responses to the advertisements they saw during the programmes. The programmes were then

divided as low, medium and high involvement based on the participants' responses. To measure the responses to the advertisements, the authors tested five aspects: viewers' recall of the ads, using unaided brand recall; aided copy point recall; copy point credibility for each message recalled; purchase interest (using a 5-point rating scale); and pre-post exposure change in brand purchase intentions. They found that in all measures, high involving programmes had the highest scores. As programme involvement increased, the ad response scores increased. For example, in terms of behavioural intentions, the mean change of viewers who had watched a low involvement programme was 6.4% compared with 14.4% among those highly involved. With purchase interest, 13.2%, the mean change among viewers with low involvement rises to 18% for viewers with high involvement.

One of their major conclusions was that programmes differed in their ability to involve the viewer. Unfortunately, although the authors tested the relationship between programme involvement and a range of ad responses, only three broad impact categories were developed: high involvement, moderate involvement and low involvement. There are, however, problems with the way each impact category was defined as being different from one another, and what programmes fit in which impact category. It was unclear how and at which point each impact category "graduated" to another. For example, it was unclear how the authors had decided what the criteria for "low involvement" was and how they were different from the criteria for "moderate involvement". This may have had to be arbitrarily determined by the researchers. In addition, it was also unclear how the three impact categories affected actual ad-viewing behaviour. The authors were more interested in post-viewing advertising responses and did not collect data on actual viewing behaviours. While they commented that viewing might be less restless and distracted in high involvement situations, they did not fully investigate that. They merely gave examples of earlier research (Leach, 1981; Rust, 1987) that supported this hypothesis; they did not validate it from their own research.

2.3.4.4 Summary

Three approaches of differentiating between programmes were presented in this section. The first approach was to classify programmes into genres of similar programme characteristics (e.g., news versus situation comedies) and to assume they have similar effects on viewers. The second approach differentiated programmes by the effort viewers put into viewing programmes (Barwise & Ehrenberg, 1988). *Demanding* programmes were liked more and found to be more rewarding to watch by viewers. These were mainly information programmes such as news, documentaries and heavy dramas. These programmes might have smaller audiences who watched them frequently and liked watching them more. Entertainment programmes were found to have less impact on viewers and they liked these programmes less as well. However, the two categories of *demanding* and *less-demanding* (entertainment) were mainly used to investigate liking and repeat viewing as an analysis of how people used TV as a medium. Barwise and Ehrenberg (1988) did not specifically investigate viewing behaviours during the programme or the advertisement break.

Another approach to differentiate between programmes was to ask viewers to rate them on pre-determined criteria. These studies attempt to measure how viewers were cognitively and affectively (e.g., Hoffman & Batra, 1991; Lloyd & Clancy, 1991) involved with the programmes. A major criticism of this approach is that the categories may have been arbitrarily and subjectively determined. It was not clear from their methodology exactly how each impact category was derived and at which point one impact category became different from the next. Some of the labels used, such as "low involvement", "moderate involvement" and "high involvement" (as used in Lloyd & Clancy, 1991), were not defined and explained fully. In particular, it was possible that some programmes belonged to more than one impact category. For example, there were programmes that were could be in both high-cognitive and high-affective

(as in Hoffman & Batra, 1991), and a separate category may be needed for these.

Next, although these studies were examining viewers' responses to advertising and the relationship between programme involvement and ad involvement, information on actual viewing behaviours was often insufficient and too generalised. Information on viewing behaviours could not be quantified as it consisted mainly of descriptions of what viewers reported they did during the ad break or how viewers had responded to the products advertised. Although ad responses were measured, information was lacking on attention and actual viewing behaviours during different programmes that involved the viewers.

The motivation for comparing the viewing behaviours with the different ways programmes have been grouped (e.g., into high or low involving programmes), is an attempt to measure not only how viewers behave in different programme situations but also whether the ways programmes have been categorised affect these outcomes. In other words, is it possible to make claims that certain programme types do involve viewers more than others. If so, what are the criteria for grouping programmes that would accurately differentiate between them?

2.4 Summary and Conclusions

The development of television audience measurement technologies was in response to rapid changes in the technological advances of television (such as digital TV) and other related forms of media such as VCR inventions, cable and satellite introductions and computer innovations. These new technologies all have an impact on the way television is used today in the same way trends in television continue to evolve with modern life. From the days where participants had to keep simple dairies, information about viewing is now recorded electronically and collected by state-of-the-art technologies that in the middle of the night "lift" information from the "black box" (actually a very small computer and modem) located in the participant's home without requiring any assistance from the participants. But this system is not without problems.

The Peoplemeter system measures mainly how many people are present during a programme and for how long. A programme's rating is mainly an indication of the size of its audience. Even so, there are suspicions that this information may be inaccurate. Although the retrieval of the information is carried out without human interference, the system is still highly reliant on participants to indicate their viewing and press the buttons (in the system) accurately enough to reflect "real" viewing. In other words, despite the large investments made in installing the system, the quality of the information may be overrated.

The Peoplemeter technology and the ratings provide quick and easily available information for advertisers. With very sophisticated smart software and a stroke on the keyboard, overnight second-by-second information on their advertisements' performance provides instantaneous results. The trouble is this information does not show how many viewers were actually "paying attention" to the advertisements.

The problem for advertisers is that the ratings for programmes could have very little to do with whether viewers are actually attentive to their advertisements. At present, advertisers pay for the cost to screen their advertisements by the ratings of the programme. Using Peoplemeter numbers and the ratings system, this would mean advertisers are also paying for "distracted viewers, rapid channel-changers, those missing from the room, and those tired of pushing the Peoplemeter buttons" (Clancy & Lloyd, 1999).

One method of knowing what is actually happening in the viewing room and to examine the quality of the viewing is to employ methods of recording audiences as they watch television. Many of these in-home observation techniques have been tested successfully, and have reported that ad viewing was accompanied by many different activities (Allen, 1965; Collett, 1986; Anderson, Field, Collins, Lorch & Nathan, 1986; Krugman, Cameron & White, 1995). An important finding is that the percentage of presence differed from the proportion of attention viewers paid to the screen. Nonetheless, although the measure of presence did not correlate with attention, there was a positive correlation between attention to programmes and attention to ads. Attention has also been known to vary with different programme types, time of day, sex and age.

It is clear from these studies that the context in which the advertisements are screened, and the type of programming that surrounds the viewing of the advertisement, both play an important role in determining the extent to which viewers watch or do not watch the advertisements. However, the precise relationship between programme characteristic, content and ad viewing is still far from clear. If involvement with programmes leads to greater attention to the advertisements, advertisers would benefit by knowing exactly which programmes these are.

One measure of involvement is the use of *eyes-on-screen* (EOS). This is the measure of the proportion of time viewers pay to the screen. This thesis proposes that it is a better measure of involvement than the measures based on Peoplemeter ratings. However, this method of obtaining a measure of "involvement" is time-consuming and not practical for everyday use. Several methods have been used previously yet none of these classification systems have examined whether viewers' actual behaviours match the predictions made for these programme classifications on the basis of viewer ratings.

One approach, developed by Barwise and Ehrenberg (1988), differentiated programmes by the level at which they make demands on the viewers. Another procedure, developed by Hoffman and Batra (1991), is based on individual personal ratings of programme impact. And a third classification is simply the system of grouping according to specific programme genres such as news, comedies and sports programmes. This method had been used by researchers who support Peoplemeter results, such as the studies carried out by Danaher (1995).

In the context of this study, the question is whether the differences in viewing behaviours for different types of programmes crossover into how people watch the advertisements embedded in these programmes. To do this, the present study first examines the proportion of time viewers spend visually attending to the screen. It focuses on how viewers behave during different programmes as a measure of involvement with advertisements. To investigate this thoroughly, viewing behaviours are then examined in relation to the three different programme classification systems discussed above.

CHAPTER THREE

RESEARCH METHOD

3.1 Introduction

The focus of this chapter describes the methodology adopted in this research project, which sought to collect "qualitative" data about the television viewing population. The approach draws on and replicates earlier in-home observational studies by videotaping viewers in their own homes. The study was conducted in Palmerston North and the data collection covered the years between 1998 and 1999.

Subsections of this chapter examine in greater detail: a brief description of the members of the sample; the recruitment method; the equipment used to record the participants; the procedure that the research took to obtain the data; the coding system and how the data were analysed.

3.2 Sample

A list of participants willing to participate in the research was first obtained by placing a request in an Omnibus survey conducted by the Department of Marketing in 1998. This is a face-to-face survey of Palmerston North households. The question asked whether respondents were willing to be contacted again to take part in a television-viewing study about the viewing behaviours of people in general. This list was then used to generate starting points for a snow-ball sample. As the researcher found someone willing to take part, the participant was then asked to refer his or her friends and neighbours who might also be willing to be contacted. As the videotaping took place, simultaneous recruitment of the panel was also taking place.

Zwaga (1992a) also used the "snow ball" technique to recruit participants in his research and mentioned the difficulties of "finding enough families to fulfil what was originally intended as a study of fifteen to twenty families" (p. 66). Some of the families who initially expressed an interest in Zwaga's (1992a) research decided not to take part after they were told that the study involved video-taping their activities for a period of 7 days. In the present study, a total of 30 families were contacted and the family who recommended the "next" family (to the researcher) were told to tell their friends and neighbours about the videotaping device. This meant that all the families who agreed to take part were fully aware that their viewing behaviours would be recorded and studied by an observer. Despite this knowledge, none of the 30 families withdrew from the study once the recording started.

At the end of the recording phase, useable taped recordings of 30 families had been collected; 16 families with children; two single member household (working professionals); eight two-member households (couples without children); one three-member household (flatmates with no children); and three four-member households (families and flatmates without children).

In order to provide statistical variability, it was originally intended that data could be collected from 30 families. But the effort to code and analyse all 30 families proved to be too difficult as the data files accumulated. The coding method required an observer to view the videotapes and a coder to input the data electronically. This meant that considerable time and financial commitments were required. As such, it was decided to investigate whether a random sample of 14 families (which had already been coded) would provide a reasonable database for the analysis. From the 14 families, approximately 350 hours of recordings were obtained. This rendered a dataset which contained the viewing of over 6000 advertisements across a range of different programmes. It was then decided that the data set would be able to provide a sound analytical base for the study.

Although there were families with children who took part in the study, this research only examined the behaviours of adult viewers. It was thought that the impact of advertising on children is a very important social issue and would warrant a separate major study on its own. There is also a strong likelihood that adult viewing behaviours would differ from children's viewing behaviours. Thus, it was hoped that this study could become a starting point for other studies using a similar methodology and coding procedure.

Zwaga (1992a) provided socio-demographic profiles of the eight families who took part in his research. In Zwaga's study (1992a), there was a brief description about each family's age and occupation as well as a drawing of each family's house-plan. In addition, the location of the television in the house and where the family normally viewed television were discussed in some detail. In the present study, demographic data such as the participants' age, occupation and household income were collected. This information was used to check the variability of the sample and to provide brief explanation of the background of the family members. However, it was not the intention of the present study to provide detailed descriptive viewing profiles of the participants, as the focus of

the analysis was to quantify incidences of ad-viewing behaviour across different programme types.

Yet, it is believed that a representative sample of a diverse background of people was recruited in this study. The sample consisted of a wide range of people of different ages, income levels and occupations. There were an almost equal number of males (45%) and females (55%). Most people (36%) in the sample were between thirty to forty years of age. Regrettably, none of the participants were older than 55 years old suggesting that older people may be under-represented in the sample. The oldest adult participants were 51 years old and the youngest 16. Seven of the participants were below the age of 30, 11 were between 31 and 40 years old, seven were 41 to 50 years old, and three were above 50 years old. Three participants did not give details of their age.

In terms of occupation, there was a good representation of people involved in professional or highly skilled jobs and in less qualified or semi-skilled employment. Professional employment included jobs in managerial positions, teaching, banking, administrators, marketing and sales, land developers, surveyors and valuers. A small number of participants (6%) were unemployed and around 10% of the participants were engaged in some sort of community service work. The majority of the participants, 23% (the largest group), were in the teaching or education profession. While 19% (the second largest group) worked in the banking industry, 16% of the participants (the third largest group) were employed in administrative positions. It could be seen from this description that people of the same occupational group tended to recommend their friends in the same industry for this research.

Most of the participants belonged to the middle to higher income group (with household annual incomes above NZ\$50,000. The exception was one participant who had an income below NZ\$20,000. She works as a part-time caregiver. There was also “a bias towards middle-class status” in Zwaga’s

(1992a) study. But this was dismissed as he argued that his study was “largely concerned with analytical generalisation” and that “representativeness is hardly the major issue at stake” (p 66). This is also the case here with regard to differences in socio-economic status. In addition, many of the participants (45%) in the present study were reluctant to divulge their household income levels. But as most of the families were friends and neighbours, one could expect the families in the sample to come from a similar social class level.

The sample in the present study consisted mainly of New Zealanders from European descent who share common cultural traditions and lifestyle patterns. It is possible that different ethnic groups use television in different ways and often this is related to their socio-economic status. For example in cases where there is poverty or lack of economic opportunities (due to ethnicity), television may be the only source of entertainment for some poorer families. Such families would therefore spend more time watching television. At least one household in the sample was from an ethnic minority group (Indian) in New Zealand. One European family also hosted a Japanese student as a boarder. In this respect, the sample contained some variation in terms of ethnic differences.

Table 2 : Summary details of participants

	No. of participants
<i>Gender:</i>	
Male	15
Female	16
<i>Age:</i>	
Below 30	7
30 to 40	11
41 to 50	7
above 50	3
Unknown	3
<i>Occupation:</i>	
Managerial	3
Teaching / lecturing	7
Administrative	5
Banking	6
Marketing & sales	2
Land development	3
Community workers	3
Home Executives	2
<i>Total household income level:</i>	
< \$50,000	1
\$50,000 to \$69,000	6
\$70,000 to \$79,000	2
\$80,000 to \$90,000	2
\$90,000+	6
Unknown	14

3.3 Equipment

To videotape the viewers, two devices known as C-Boxes (first developed by Collett in 1986) were constructed. Each C-Box consisted of a 25-inch colour television, a cabinet that housed up to four video recorders and a hidden, outwards facing, video camera. The television sat on top of the cabinet, above one of the video recorders that was available for the use of the participants. The other three recorders were placed on top of each other in the locked part of the cabinet to prevent the household members from interfering with them. These were used to record viewers' activities while they watched TV.

The three research video recorders were rewired and programmed, to record continuously, so that when one tape ended, it triggered off the recording in the next recorder. As each recorder was capable of recording for a maximum of 8 hours, continuous recording of 24 hours was possible before the tapes needed to be changed. The researcher therefore did not have to visit the family too frequently to change the tapes. This was an attempt to minimize participants feeling they were being videotaped and therefore remain conscious of their behaviour in front of the television set. The cabinet also concealed a small video camera installed behind a piece of tinted glass placed across the top shelf. This hidden camera recorded the activities in the room, particularly those of people in front of the TV set.

A time and date generator was also wired to the recorders so that time and date details were superimposed on the recorded tapes. This generator was first programmed to set the starting time and date of the recording for each family. Thereafter, every time the television was switched on, the generator automatically recorded the time and date on to the tapes. Coders later used the date and time details on the tapes to identify the changes in the behaviour of the viewers for each programme and advertisement. The video recorders were also re-engineered to include a picture-in-picture facility from the TV so that what

was being screened on television at the time of the recording was also included in the tapes. When the recorded tapes were played, this information appeared as an inset on the lower right hand side of the frame. This enabled the researcher to know what programme and advertisements were being screened during the taping.

It was thought perhaps some viewers might encounter occasions when they wanted to stop the taping, for example when guests visited the families and might prefer not to be in the study. To stop the recording momentarily, there was a small white button at the back of the cabinet wired to control the recorders. When pressed, this would discontinue the recording. The recorders reset, and taping resumed as soon as the television was switched on again.

Within the range of the camera lens, each C-Box recorded the viewers and what they were doing. The set was wired in such a way that as soon as the television was turned on, the recording started. When the television was switched off, recording continued for 30 seconds.

3.4 Programme Involvement Rating Scale

A rating scale consisting of 18 questions was devised to measure respondents' involvement with different programmes. This scale was adapted from a combination of Hoffman and Batra's (1991) and Lloyd and Clancy's (1991) questions used in their respective studies. A copy of the scale is provided in Appendix B.

These rating scales were used as they have already been well tested in other research contexts. The Hoffman and Batra (1991) scales were used to classify the recorded programmes as high or low impact. According to Hoffman and Batra (1991), the questions from their scales were able to draw out both cognitive and affective responses from the viewers. They measured the viewers' involvement in terms of their intentions to watch, appreciation for and the impact of the programme. The scale included items that asked viewers how they felt about the programmes, whether they had learned anything, and the amount of effort they had invested watching the programmes. The respondents graded the programmes in terms of whether they thought the content and programme delivery were successful and appealing, how their emotions were aroused and whether they wanted to get more from the viewing experience.

The questions adapted from Lloyd and Clancy (1991) evaluated the value of the programme in terms of how the programme was able to "entertain" the viewers. According to them, the success of this entertainment function would bring about higher levels of viewer involvement with the programme. The questions included items that measured the viewer's personal empathy with the main characters of the programmes and how the programme's content created enjoyment and personal meaningful experiences for viewers.

3.5 Procedure

Recordings of the families started in October 1998 and continued until December 1999. The C-Box and TV set were placed in each participants' homes over a period of around 8 days.

Household members were given an information sheet to read that provided information about the equipment, the contact details of the researcher, the purpose of the investigation and the process of the research. In line with the New Zealand Market Research Society Code of Ethics, participants were assured the information they provided would be treated as confidential and they could withdraw from the study at anytime. They could also have any part of the tapes erased (if they so wished), and had the right not to agree to the tapes being archived by the researchers. They were also given the assurance that the information on the tapes was to be used only for the purposes of the study or teaching and that the tapes would only be viewed by the researcher and her assistants. A consent form was presented to the family where all adult participants (over 16 years old) gave their agreement. A copy of the information sheet and consent form are provided in Appendix C and Appendix D.

The researcher visited the household the day following the installation to check the equipment as well as to judge whether the tapes needed to be changed. This was also to estimate the amount of viewing the family did in 1 day and then use that to estimate the number of days needed before the tapes "ran out". The researcher then, upon her discretion, visited the family again to collect and to change the tapes.

Within 1 week after the removal of the cabinet from the household, the researcher visited the family again. In interviews with family members, some demographic details were recorded. But the main purpose of the visit was to

obtain, using the 18-item scale, the participants' personal evaluation of the programmes they had watched during the research period.

3.6 Coding

The coding process consisted of two parts. In the first part of the process, the researcher coded information about the household's overall viewing over 8 days. This coding was similar to a visual diary of what programmes and advertisements the family members had watched, the channels to which they had tuned, who was "watching", and the lengths of time viewers were present in the room. The second part of the coding process included a more detailed log of what viewers were observed to be doing during the ad breaks.

3.6.1 Coding overall viewing behaviours

The process of viewing the tapes took a long time as the tapes were recorded in real time (not time-lapsed). As soon as each tape was filled, the researcher collected it and the data were coded. The tapes were first "watched" to identify any irregularities of the taping and to record the family's overall viewing. Information on who was watching, what channel was being watched and what programme and advertisements were screened was manually coded.

Initially, each tape was viewed to identify the names of the programmes watched by the household. Then observational data of the viewing were transposed onto a record sheet that registered the date, time and name of the television programme watched, its channel, the number of people and who was in the room. Each viewer in the household (including children and guests) was given a coded number. Time and date details of the following events were entered onto the log sheets:

- every time the channel was changed

- every time the television material changed (e.g., programmes to commercials)
- every time someone entered or left the room.

For programmes, start time was measured from the beginning of the opening music sequence. Likewise, the finish time was recorded after the music and credits had ended. Where possible, the name of the programme was identified and recorded (using published television schedules). If not possible, it was treated as an "unknown programme". Advertisements were also recorded on a separate column.

When it was possible to know which position the ad pod had appeared in within the programme, it was recorded. The first ad pod was taken as the first set of advertisements that appeared after the starting time of the programme being screened, the second position ad pod was the second set of advertisements to be screened, and so on. Within each ad pod, the position of each advertisement within each individual ad break was also recorded.

To differentiate between advertisements, the brand of the product or the name of the advertiser was identified and coded. In addition, other information such as announcements, trailers, promotions for up-coming programmes for the station, music interludes and so on were also included in the coding sheet.

In coding the members of the family, "P" stands for "Person". Odd numbers were used for males, and even numbers were used to distinguish females: "P1" was the oldest male participant in the family, "P2" was the oldest female; the next oldest male was "P3", and the next oldest female was "P4".

3.6.2 Ad pod selection

As one of the objectives of this study was to describe the relationship between programmes and ad viewing behaviours, information on viewing of programmes preceding the ad pods was necessary. To determine whether the ad pods analysed were relevant, it was important to consider whether viewers were actually watching the programme preceding the advertisements. There would be little justification for studying the effect of the programme on ad-viewing behaviour if the person was only there for a few seconds during the programme directly preceding the ad pod. This would show that the viewer was not directly "involved" with the programme and would devalue any assessment of ad viewing behaviours displayed during the programme.

This problem was also evident to earlier researchers interested in the relationship between programme involvement and ad involvement. After evaluating the assumptions used by earlier research, three criteria that would show there was involvement during the programme preceding the ad break (analysed) were developed for this study. In comparisons between programme behaviour and commercials behaviour, Krugman et al. (1995) argued that because the average commercial break time (in the United States) was about 3 minutes long, the data for programme viewing behaviour should be comparable. Therefore, they had observed viewers watching a certain programme for 3 minutes. The average commercial break in New Zealand is around 4 minutes. As a starting point, it was decided to use this 4-minute criterion in the same way. This was the first criterion used in this study in selecting ad pods to be analysed: that the person must be in front of the TV watching the programme 4 minutes before the start of the commercial break.

But being there at the 4-minute "cut-off" time did not mean that the person was there all of the 4 minutes. Danaher and Lawrie (1998), suggested the amount of time a person spent watching a programme showed the commitment the person

had for the programme. To be a committed viewer, he or she must have watched at least 80% or more of an entire programme. A second criterion was therefore developed in the present study, that a viewer must be there for at least 80% out of the 4 minutes preceding the ad pod. This would show the person was somewhat "committed" to the programme. In another study, Danaher (1992) selected the channel (i.e. the programme) to which viewers were tuned for the majority of the 30 seconds before the ad break (p. 41).

A third criterion developed for this study was to consider only continuous viewing before the ad break. For this, it was decided that a viewer must be continuously watching for at least 30 seconds preceding the commercial break. This method was also used by Danaher (1995) to determine the pre-ad-break channel to which a person was tuned to before an ad break.

Clearly, there were a number of criteria that could be used to select the relevant sample of ad pods to be analysed. The important issue was whether the use of one criterion would result in a different outcome from another criterion used. A comparison of the data obtained using the three sets of criteria, (i.e. present 4 minutes before ad pod; present for at least 80% of the 4 minutes before the ad pod, and watched at least 30 seconds continuously before ad pod), showed negligible differences in the results (see Appendix F for details). For this reason, the following analysis uses the criterion that the viewer must be continuously watching 30 seconds before the start of the ad pod.

3.6.3 Behaviours during ad breaks

For each selected ad pod, the behaviour of each viewer was coded on coding sheets at 3-second intervals. Thus a 30-second advertisement resulted in 10 separate observations of behaviour for each person appearing in the tape. A 2-minute long advertisement generated a total of 40 observations, and so on.

For each observation, the researcher noted whether the viewer's eyes were oriented towards the screen (EOS) and whether the viewing was accompanied by another activity.

The activities that accompany looking at the screen may include activities such as eating and drinking, doing household chores, sewing, feeding children, stroking a pet or smoking. These activities have been noted as "complementary" activities by Krugman, Cameron and White (1995). Other activities, such as attending to others (e.g., children or pets), playing games, muting the sound, reading and writing, sleeping, taking part in conversations, and doing things around the room, are known to be "competing" for attention with the set. The complete list of coded activities is provided in Appendix G.

3.6.4 Programme codes

All the ad viewing observations occurred within programme environments. In television research, different researchers with different research objectives had used different ways to group programmes. In this study it was decided to make comparisons of viewing behaviours using the three programme classification systems below:

- the way programmes make "demands" on viewers
- common genre type categories, such as news, comedies and sports programmes
- the impact viewers reported the programme had on them.

The first category of programmes was coded following the classification developed by Barwise and Ehrenberg (1988). Programmes such as news and information programmes (documentaries and current affairs) were grouped as *demanding* programmes. Other programmes such as light entertainment, films,

light drama and sports were categorised as *less-demanding*. Within the light entertainment category, programmes included comedies, variety and music shows, games and chat shows. Movies were included in the category of films. Light drama included mini-series, situation comedies, plays and weekly dramas. Sports were actual sporting events that were broadcast. Children's programmes were excluded from the classification.

The second classification system, that of grouping by common genre characteristics, was adapted from Danaher (1995). In grouping the programmes into genre categories, 11 programme types were developed for this study. They were chat shows, comedies, game shows, news / documentaries, drama, informative interest shows, musicals, soaps and mini-series, competitive entertainment shows, sport, and movies. As only a small number of children's programmes were watched, they were coded but excluded in the final analysis.

For the third category, programmes were grouped into how they had impacted on the viewers, based on the involvement rating scales administered during post-research interviews. Following the example of Hoffman and Batra (1991), each individual programme was assigned to an impact type category.

Hoffman and Batra (1991) developed three clear categories of programme impact: High Affective, High Cognitive and Low Impact. Three variables were used as measures of these impact dimensions: "FEELING", "GET MORE" and "LEARN". FEELING measured the feelings evoked by the programme. GET MORE measured whether viewers watched the programme to get more out of the experience or merely to pass the time. LEARN measured whether viewers learned anything from the programme.

To reduce the difficulties in the measurement, only two variables of the three original variables in the Hoffman and Batra study were included in the present study. It was thought the variable defining FEELING correctly measured the

affective aspect of involvement and was used. The other two variables, GET MORE and LEARN had very close co-efficient results in Hoffman and Batra (1991). Therefore, it was decided that one variable would sufficiently represent this dimension. The variable LEARN was chosen because it was thought that it was a more accurate measurement of impact than GET MORE. The latter variable was perhaps more relevant in measuring the intention of watching the programme (as an indirect indication of involvement) rather than in assessing its direct impact on viewers. LEARN, on the other hand, implied viewers were cognitively involved if they had learnt from the programme. For this study, on the basis of the two variables, FEELING and LEARN, programme rating scores were obtained.

Hoffman and Batra (1991) maintain that some programmes have the ability to involve both cognitively and affectively. According to them, High Affective and High Cognitive programmes are High Impact programmes. They did not devise a separate category to represent these programmes. The present study, however, sought to group programmes clearly by all the ways in which the two impact types, affective and cognitive, could have affected the viewers. This resulted in the following four categories:

High Cognitive : where viewers rated the programmes as those that did not touch their feelings but which they had learnt from

High Affective : where viewers rated the programmes as those that touched their feelings but which they did not learn from

Low Impact : where viewers rated the programmes as those that did not touch their feelings and which they did not learn from

High impact : where viewers rated the programmes as those that both touched their feelings and which they also learnt from

Due to different methodologies used, the coding of the viewing behaviours in the present study differed from that of Hoffman and Batra (1991). For instance, the activity of "talking during the ad" was analysed in relation to whether this activity accompanied visual attention. Therefore, when viewers were talking and had eyes-on-screen, this activity would be analysed as "*EOS+behaviour*". If there was talking but no visual attention, this would have been classed as "*No EOS*".

3.7 Analysis

3.7.1 Complete ad pods

Before the coding could be carried out, a preliminary decision had to be made. The amount of information on the tapes recorded was formidable, consisting of: different forms of viewing patterns; lengths of viewing times; and interrupted commercial viewing (due to switching channels and turning off the television set). It was therefore decided that only continuous commercial breaks (full ad pods) were to be analysed. These continuous ad pods were "complete" in the sense that each ad pod contained the full set of advertisements screened during the ad break. While this introduces a potential bias in the results, this bias is likely to be positive in the sense that incomplete viewing might suggest a response to television advertising. Other incomplete ad pods could not be analysed since there was incomplete information recorded on the tapes about the advertisements. Using the criterion of the complete ad pods, it was possible to work out the types of advertisements screened, how long each advertisement lasted, and the number of advertisements in each ad pod. It was then possible to analyse the proportion of time during which viewers were attending to the advertisements and, if they were not, how long they spent on other activities during the complete ad break. This analysis was not possible for incomplete pods.

There were a total of 869 complete ad pods in the initial analysis. Each ad pod ranged from a minimum length of around 2 minutes to a maximum length of around 5 minutes.

Viewing behaviours were analysed in terms of the degrees of visual attention paid to the screen during the complete ad pods using four categories: "full attention" to the screen, coded as *EOS only*; "partial attention" when viewers were watching but concurrently doing something else (coded as *EOS +*

behaviour); "no attention" to the screen (coded as *No EOS*); and left or not in the room (coded as "not present"). The categories were all calculated in the same way. The data used in the analysis is the proportion of time spent in one of the categories. A measure of "presence" (measured as the sum of the scores of *EOS only*, *EOS + behaviour* and *No EOS*) was also computed.

These four viewing behaviour categories were used to examine the different programme classification systems of Danaher (1995), Barwise and Ehrenberg (1988) and Hoffman and Batra (1991).

While it was possible to analyse the viewing behaviours directly for programme types using Barwise and Ehrenberg (1988) and Danaher (1995), a slight modification was made to the categories proposed by Hoffman and Batra (1991). The analysis included examining viewing behaviours in an additional impact category not included in Hoffman and Batra (1991). As a consequence, it is not possible to make a direct comparison with Hoffman and Batra's (1991) findings on viewing behaviours during programmes.

CHAPTER FOUR

RESULTS

4.1 Introduction

The previous chapter outlined the major steps taken to collect data for this research, and described briefly the participants who agreed to take part and the equipment used to videotape them. It also elaborated how the taping, the coding and the analysis of the data were carried out. This chapter reports the results of the study. It gives details of the proportion of time viewers spent watching television, and how much attention they paid to advertisements.

The first part of the chapter describes the overall viewing patterns of the families, including the amount of time each member of the family was seen in front of the TV; the types of programme most often watched (by each person) and the amount of attention paid to advertisements in general. Subsequent sections of this chapter report on ad-viewing behaviour organised by three programme type classification systems used by other researchers described in section 2.3.4.

The results reported in this chapter show the effects programmes have on viewers' attention to the advertisements and whether these behaviours differed when different programme classification systems were used.

4.2 Overall Viewing Patterns of the Participants

4.2.1 Viewing time

Table 3 shows a breakdown of household television viewing time for each family. It shows the total time the TV was on, the total time when there was no one in the room, and the total time the household spent in front of the TV. Table 4 presents data on the amount of time each member of the family was observed as being present when the TV was on, either by themselves or with other family members.

Table 3 is sorted by total household viewing.

Table 3: Household television time

	Days TV set in house	Time TV set turned on	Time no one in view	Total household viewing
Family 11	8 days	36:19	2:50	33:29
Family 13	8 days	40:59	8:54	32:05
Family 1	8 days	34:09	5:02	29:07
Family 2	8 days	28:52	3:09	25:43
Family 6	8 days	28:55	5:43	23:12
Family 12	8 days	26:09	3:03	23:06
Family 10	8 days	23:02	1:36	21:26
Family 3	8 days	26:52	6:48	20:05
Family 14	8 days	21:17	2:41	18:35
Family 4	8 days	18:12	0:36	17:36
Family 8	8 days	24:21	7:12	17:10
Family 7	8 days	17:08	0:32	16:36
Family 5	8 days	15:57	2:35	13:23
Family 9	8 days	11:49	2:11	9:38
Average		25:17	3:47	21:31

The amount of time the television set was on varied from family to family, ranging from approximately 12 hours to 41 hours. On average during the 8-day observation period, over all the households, the average time the television set was switched on amounted to slightly above 25 hours.

Not all viewers were always present when the TV was on. Again there are variations across the families. The time when the TV was playing to no one, ranged from as much as 9 hours to as little as 30 minutes. The average time where no one was around during the screening was just under 4 hours (3:46).

Table 4: Time individual household members were present

	Presence when TV set was on			
	P1 in view	P2 in view	P3 in view	P4 in view
Family 11	8:35	9:00	-	-
Family 13	11:20	5:58	-	-
Family 1	10:54	13:38	-	-
Family 2	13:11	10:30	-	-
Family 6	19:18	16:37	-	-
Family 12	10:57	10:26	-	-
Family 10		2:26	-	-
Family 3	8:34	6:23	-	-
Family 14	12:05	7:57	-	-
Family 4	6:39	0:21	9:56	6:45
Family 8	7:18	9:13	10:11	-
Family 7	5:57	9:10	10:39	-
Family 5		7:46	-	10:05
Family 9	7:20	6:43	-	-
Average	10:10	8:20	10:15	6:45

P1:Oldest male P2:Oldest female P3:Next oldest male P4:Next oldest female

Note: The rows of this table do not add to the row total in Table 3 because children's viewing is excluded and because of duplicated viewing.

As shown in Table 4, there are small differences in the proportion of time each family member was seen in front of the TV. During the week, individual members of the family tended to spend similar amount of time "watching" TV.

4.2.2 Types of programmes

Overall, viewers spent more of their time watching News than other programmes. These programmes accounted for 15% of the time viewers spent in front of the TV. Table 5 shows the percentage of time family members were present during different programmes.

Table 5: Proportion of time viewers were present during different programmes

Programme Type	%
News	15
Childrens' programmes	11
Comedy	10
Soap	9
Movies	8
Current affairs	6
Popular Drama	5
Sports	4
Documentaries	3
Story dramas	3
Reality TV	3
Musicals	3
Quiz & game shows	2
Chat shows	2
Sci-fi	2
Personal interest	2
Lifestyle /Home	1
Police True Stories	1
Variety Entertainment	1
Unclassifiable programmes	9
Total	100

As most of the family members tended to watch TV together as a family, there were many occasions where adults were present during children's programmes (11%). Other types of programmes most "watched" were comedies (10%), soaps (9%) and movies (8%).

4.3 Behaviour during Advertisement Breaks

This section reports the viewing behaviour observed during the advertisement breaks. These behaviours were examined in four ways: eyes-on-screen only (*EOS only*), meaning the person was seen looking at the screen; eyes-on-screen and engaging in one or more simultaneous activities (*EOS + behaviour*); not looking at the screen (*No EOS*); and not present. Table 6 shows the results of the four viewing behaviours in which viewers were observed to be engaged during the complete ad pods.

Overall, viewers were not present for 23% of the time complete ad pods were playing. Viewers spent around 14% of the time with their eyes-on-screen only and around 11% of the time they were looking while at the same time doing something else. Just over half the time (52%), they were not looking at the screen at all (See Table 6).

Table 6: Viewing behaviours during complete ad pods

No. of ads observed	<i>Present</i>			<i>Not present</i>	<i>Total</i>
	<i>EOS</i>	<i>EOS + behaviour</i>	<i>No EOS</i>		
	%	%	%	%	%
6017	14	11	52	23	100

Note: Viewer must be present for 30 seconds before the ad breaks.

In order to make possible comparisons with Peoplemeter data and viewing behaviours, an index of presence was developed. The previous section analysed ad-viewing behaviours in a single variate way (i.e., *EOS only*, *EOS + behaviours*, *No EOS* and *Not present*).

A number of new variables were computed, adding *EOS only*, *EOS + behaviour* and *No EOS*, to give the proportion of time in the room or “presence”. *Total EOS* shows the proportion of time when there was visual attention regardless of whether it was the sole activity or not. Furthermore, computing the ratio of *EOS only:Presence* or *Total EOS:Presence* gives indices of ad-viewing quality. For example, *EOS only:Presence* shows the proportion of time when viewers had their eyes-on-screen only during the time viewers were present. On the other hand, *Total EOS:Presence* shows the proportion of time viewers had their eyes-on-screen including those times when they were also engaged in other activities.

The two ratios enable comparisons to be made between visual attention to the ads and "presence". They provide information on how viewers pay "attention" to the advertisements during the time the viewers were present.

Table 7 shows details of the total attention paid to the screen, and the total amount of time viewers were present during the ad breaks. In addition, it also

reports the time viewers paid to the screen as a proportion of the time they were present.

Table 7: Attention during complete ad pods

No. of ads observed	<i>EOS only</i>	<i>Total EOS</i>	<i>Presence</i>	<i>EOS only</i>	<i>Total</i>
		(EOS + EOS + behaviour)	(EOS + EOS + behaviour + No EOS)	:	<i>EOS :</i> Presence
	%	%	%		
6017	14	25	77	18	32

Note: Viewer must be present for 30 seconds before the ad breaks.

Viewers spent only 25% of the time the ads were on, engaged in some sort of eyes-on-screen (*Total EOS*), which was just slightly greater than the time spent not present at all (23%). Even when present, viewers spent only 18% of the time eyes-on-screen (*EOS only:Presence*) and only 32% of the time engaged in some eyes-on-screen attention (*Total EOS:Presence*). Thus, for the time viewers were present during the ad breaks, only for around one-third of the time, was some sort of attention paid to the screen. The figure was even lower when considering full eyes-on-screen attention.

4.4 The Relationship between Programme Involvement and Ad Involvement

A major aim in this study was to investigate whether the proportion of time viewers paid attention to television advertising was affected by how programmes influence the viewer. To investigate the relationship between programme involvement and ad involvement, ad-viewing behaviour was observed over a large range of different programmes. These programmes were grouped in three different ways, using three programme classification systems. The three ways were developed or used in the research by Barwise and Ehrenberg (1988), Danaher (1995) and Hoffman and Batra (1991).

For each programme classification system namely those developed by Barwise and Ehrenberg (1988), Danaher (1995) and Hoffman and Batra (1991), a series of chi-square tests was performed to find out whether the programmes differed in terms of ad-viewing behaviour. These tests were carried out to determine whether the programme categories were valid in predicting ad-viewing behaviour. If the behaviour during ads differs for different programme classifications, then the programme classification could be used as a surrogate for behaviours during the ads within those programmes.

4.5 Ad-break Behaviour by Programme Type Using the Classifications Developed in Barwise and Ehrenberg (1988)

Barwise and Ehrenberg (1988) suggested that the viewer as "an investor of effort" would put more effort into watching programmes that had a greater impact on them. They classified informational programmes as *demanding* programmes, and entertainment programmes, which viewers watched mainly for relaxation, as *less-demanding* programmes. While Barwise and Ehrenberg (1988) reported on general television viewing behaviour, the present study used these two categories to test whether there was a relationship between "investment of effort" in watching a programme (using the Barwise and Ehrenberg (1988) classifications) and the attention paid to ads. One possibility is that one might expect that more attention would be paid to ads shown during *demanding* programmes. Conversely, *demanding* programmes might be associated with less attention to the ads simply because viewers need a break.

Table 8 shows the viewing behaviours broken down by the two sets of programme categories used in Barwise and Ehrenberg (1988). Informational programmes such as news, documentaries and current affairs programmes were included as *demanding* programmes while all other programmes that mainly provided entertainment value were grouped as *less-demanding*.

The results reported in this section include all programmes recorded in this study except for children's programmes (which was not part of Barwise and Ehrenberg's (1988) original list of programme types).

4.5.1 Comparison of behaviour during *demanding* and *less-demanding* programmes

According to Barwise and Ehrenberg (1988), there were marked differences between "*demanding*" information programmes and "*less-demanding*" entertainment programmes in terms of both claimed repeat viewing and liking for programmes. They suggest that the characteristics of these two programme types differ in the way they make the viewers "think" or "help them relax". Ehrenberg and Goodhardt (1981) suggest that viewers generally regard *demanding* programmes as more involving than the *less-demanding* ones. However, in the present study, in terms of actual viewing behaviours observed during the ad break, the results did not indicate that the differentiation between the two programme type classifications was useful when applied to ad-watching behaviours (See Table 8). The difference in the averages between ad-viewing behaviour for *demanding* and *less-demanding* programme types was not significant ($\chi^2 = 5.5677$, $df = 3$, $p > 0.05$)

Table 8: Ad-break behaviours by programme type using Barwise and Ehrenberg (1988) categories

Programme type	No. of ads observed	<i>Present</i>			<i>Not present</i>	Total %
		%			%	
		<i>EOS only</i>	<i>EOS + behaviour</i>	<i>No EOS</i>		
<i>Demanding</i>						
Information	1067	17	15	49	19	100
News	1135	15	10	41	34	100
Average		16	13	45	26	100
<i>Less-demanding</i>						
Light Ent	619	14	5	54	27	100
Light drama	2122	14	10	60	16	100
Films	752	4	6	67	23	100
Sports	231	4	20	60	16	100
Average		9	10	60	21	100
Overall average		11	11	55	23	100

Note: Viewer must be present for 30 seconds before the ad breaks. % is subject to rounding error.

4.5.2 Comparison of programmes within *demanding* programmes

Both types of demanding programmes demonstrated a similar pattern. Within the category of *demanding* programmes, the difference between news and information type programmes was not significant ($\chi^2 = 6.3747$, $df = 3$, $p > 0.05$). On average, viewers tended to be absent for about a quarter of the time (26%)

and were not looking for about half the time (45%). Just over a quarter of the time (26%) was spent attending to the screen. For both types of *demanding* programmes, when viewers were looking at the screen they tended to spend more time *EOS only* than EOS while concurrently doing something else.

4.5.3 Comparison of programmes within *less-demanding* programmes

There were significant differences between the four programmes in the *less-demanding* programmes category ($\chi^2 = 31.3000$, $df = 9$, $p < 0.01$). Thus, grouping these programmes in a single category is not appropriate. The behaviours during the ads for these four programme types were different. For example, light entertainment and light drama programme types had much higher *EOS only* scores (14%) than films and sports (4%). In other words, people tended to spend more time in full EOS for light entertainment programmes (14%) and spent less time engaged in other activities (*EOS + behaviour* = 5%). In comparison, during sports programmes, there was less full attention (4%) and more time spent doing something else during the ad viewing (*EOS + behaviour* = 20%).

In terms of similarities between these four programme types, films and sports programmes seemed to be different from light entertainment and light drama. Light entertainment and light drama programmes appeared to be more like *demanding* information and news programmes. In other words, within the range of entertainment programmes, some were more alike than others.

4.5.4 Comparison of visual attention levels during *demanding* and *less-demanding* programmes

According to Barwise, Ehrenberg and Goodhardt (1979), audience size did not influence viewers' appreciation of programmes. They reported that the audience size for information programmes (*demanding*) was on average smaller than for entertainment programmes (*less-demanding*). But in spite of their smaller audiences, *demanding* programmes had higher appreciation scores on average.

Table 9: Ratios of visual attention to actual presence

Programme type	No. of ads observed	Total EOS (EOS only +		Presence (EOS only +		Total EOS : Presence %
		EOS only %	EOS + behaviour %	EOS + No EOS) %	EOS only : Presence %	
<i>Demanding</i>						
Information	1067	17	32	81	21	40
News	1135	15	25	66	23	37
Average		16	29	74	22	39
<i>Less-demanding</i>						
Light Ent	619	14	19	73	19	26
Light Drama	2122	14	24	84	17	29
Films	752	4	10	77	5	13
Sports	231	4	24	84	5	29
Average		9	19	80	11	24
Overall average		11	22	78	15	29

Note: Viewer must be present 30 seconds before the ad breaks. % is subject to rounding error.

χ^2 for EOS / Presence was calculated by using EOS only % by presence % minus EOS only %

χ^2 for Total EOS / Presence was calculated by using Total EOS % by presence % minus total EOS

A similar observation can be made in the present study. It is apparent from Table 9 that on average, ad breaks during *demanding* programmes had fewer viewers than those in *less-demanding* programmes (Presence = 74% versus 80%). But the proportion of time full attention was paid (*EOS only*) was higher for *demanding* than for *less-demanding* programmes (22% versus 11%). Thus viewers spent more time watching the ads during demanding programmes even though less time was spent present during the ads for these programmes.

There was a close to significant difference between *demanding* and *less-demanding* programmes ($\chi^2 = 3.0409$, $df = 1$, $p = 0.0812$) at the 5% level for the ratio, *EOS only*:Presence. Also, on average, the ratio of *Total EOS*:Presence, was also significantly higher for the more *demanding* programmes than *less-demanding* programmes (39% versus 24%) ($\chi^2 = 4.2711$, $df = 1$, $p < 0.05$). Results in the present study indicate that the ads screened during *demanding* programmes were watched by a smaller audience but from those present, there were generally higher levels of attention paid to them than to *less-demanding* programmes.

4.5.5 Comparing viewing behaviour within *demanding* programmes

It can also be seen from Table 9 that ad-viewing behaviour during both news and information programmes were similar. Neither *EOS only* nor *Total EOS* within *demanding* programmes differed significantly ($\chi^2 = 0.0646$, $df = 1$, $p > 0.05$, and $\chi^2 = 0.0406$, $p > 0.05$). Thus it makes sense to combine the two programmes into a single category.

4.5.6 Comparing viewing behaviour within *less-demanding* programmes

In analysing visual attention behaviour in time present during the ad break for *less-demanding* programmes (see Table 9), there was significant difference in terms of the ratio of *EOS only* :Presence ($\chi^2 = 13.3583$, $df = 3$, $p < 0.05$) but not so in terms of *Total EOS*. The statistic for the ratio of *Total EOS*:Presence, however, was just slightly greater than 0.05 ($\chi^2 = 7.1580$, $df = 1$, $p = 0.0670$). This means there were differences between the two categories of programme types within the *less-demanding* programmes only in terms of the proportion of time viewers were looking fully at the screen (*EOS only*). These results suggest it is not sensible to combine these programmes when considering viewing behaviour.

According to Barwise, Ehrenberg and Goodhardt (1979), within the categories of *demanding* and *less-demanding* programmes, the correlation between ad ratings and appreciation scores was positive. Higher rating shows tended also to be given higher appreciation scores by their viewers. The present study did not measure appreciation scores and cannot make direct comparisons, but it is possible to compare viewing behaviour with presence. Ad-viewing behaviour and ad ratings for *less-demanding* programmes appear to have an inverse relationship. For example, Sports shows had the highest proportion of viewers present (84%) of the four *less-demanding* programme types, but the proportion of time where ad viewers had their eyes fully on the screen was lowest (5%). In contrast, Light entertainment programmes, which had the lowest proportion of time where viewers were present (73%), had the highest EOS score (19%).

4.5.7 Summary

The differences in the programme types as per Barwise and Ehrenberg's (1988) classification were sometimes significant and sometimes not, depending on which index was used. There was no significant difference between the two programme categories by viewing behaviours ($\chi^2 = 5.5677$, $df = 3$, $p > 0.05$). However, by the ratio of *Total EOS:Presence*, there was significant difference at the 5% level ($\chi^2 = 4.2711$, $df = 1$, $p < 0.05$). By the ratio of *EOS only:Presence*, the difference between the two programme types was slightly higher than the 5% confidence level ($\chi^2 = 3.0409$, $df = 1$, $p = 0.08$).

Within *demanding programmes*, information programmes and news were closely associated in terms of visual attention: both had high *EOS only* and *Total EOS* scores. But their difference was not significant. In contrast, within the *less-demanding* category, different programmes varied differently in terms of attention paid to the screen and presence in the room. And the difference between the *less-demanding* programmes was significant. Moreover, some *less-demanding* programmes appeared to be similar to demanding programmes. Using *EOS only* scores, light entertainment and light drama (*less-demanding*) resembled information and news (*demanding*) programmes.

In conclusion, this study found no significant difference between *demanding* and *less-demanding* programmes, when the overall viewing patterns (including not present) were considered, but differences were almost significant at the 5% level by *EOS only* and significant for *Total EOS*, when behaviour during presence was considered. Thus there is some support for Barwise and Ehrenberg's (1988) distinctions of these two categories of programme types. Both types of *demanding* programmes demonstrated a similar pattern (no significant difference between them) and could be categorised as a single category. But the non-significant difference between the *less-demanding* programmes suggests they should not be classified together. Furthermore, an

inspection of the viewing behaviour patterns showed some of the so-called *less-demanding* programmes had patterns very similar to those in the *demanding* category. Thus the two categories may apply for measures of liking and repeat viewing of programmes (proposed by Barwise and Ehrenberg, 1988), but the results reported here suggest that the categories are not appropriate for advertising behaviour.

4.6 Ad-break Behaviour by Genres

In the previous section, different programmes were first grouped into genres and classified as *demanding* or *less-demanding*. This section examines viewing behaviours for programmes categorised into broad genre types commonly used in the television industry. Seven of the eleven programme types listed in this table were replicated from Danaher (1995). Danaher's (1995) study was concerned only with prime time programmes, so did not present data for non-prime time programme types. In the present study, it was necessary to develop additional categories to include non-prime time viewing. The four categories not found in Danaher (1995) were adapted from the NZOnAir (2000) report on local television. These programmes were chat shows, informative interest shows, competitive entertainment shows, and musicals. The results for the four viewing behaviour categories (*EOS only*, *EOS + behaviour*, *No EOS* and *Not present*) are shown in Table 10.

The data presented in Table 10 were ordered by percentage of time spent *EOS only*.

Table 10: Viewing behaviour by genre categories

Programme types	No. of ads observed	Present		Not present		Total
		%		%		%
		<i>EOS only</i>	<i>EOS + behaviour</i>	<i>No EOS</i>		
Chat shows*	138	35	8	27	30	100
Comedy	524	21	13	50	16	100
Game shows	209	20	6	60	14	100
News/Docu/Current Aff	1984	17	12	45	26	100
Drama	968	17	10	60	13	100
Informative Interest*	218	12	17	61	10	100
Musical*	202	7	10	54	29	100
Soap/mini series	630	7	6	61	26	100
Competitive entertain*	70	5	59	24	12	100
Sports	231	4	20	60	16	100
Movies	752	4	6	67	23	100
Average		14	15	52	20	100

Note. Viewer must be present for 30 seconds before the ad breaks. % is subject to rounding error.

* refers to programmes not originally found in Danaher (1995)

It is apparent from Table 10 that, as might be suspected from the results presented earlier, programme types varied consistently in terms of the four viewing behaviours ($\chi^2 = 366.28112$, $df = 33$, $p < 0.01$) particularly with regard to the intensity of the visual attention paid to them (i.e. *EOS only* or *EOS + behaviour*).

The spread of the variation of the programme was rather wide, regardless of which visual attention measure (*EOS only* or *EOS + behaviour*) was used. For example, the variation between the programmes, for the *EOS only* score, was between 4% and 35% (a six-fold difference). The scores for *EOS + behaviour* range between 6% and 59% (a ten-fold difference). For *No EOS*, the range was from 24% and 67% (a three-fold difference). In addition, the ordering of the genres differs according to the attention measures. These results show viewers paid varying degrees of attention to the different programmes, and these differences between the programme types were quite large.

4.6.1 Quality of viewing across different genres

This section presents results on the quality of viewing across different programmes, while present. The analysis examines particularly the relationship between attention paid to the screen (as measured by *EOS only* and *Total EOS*) and presence.

Table 11: Visual attention and presence

Programme types	No. of ads observed	Presence				
		<i>EOS only</i>	<i>EOS +behaviour</i>	<i>EOS +behaviour + No EOS</i>	<i>EOS only</i>	<i>Total EOS</i>
		%	%	%	%	%
Chat shows	138	35	43	70	50	61
Comedy	524	21	34	84	25	40
Game shows	209	20	26	86	24	30
News/Docu/Current Aff	1984	17	29	74	23	39
Drama	968	17	27	87	20	31
Informative Interest	218	12	29	90	13	32
Musical	202	7	17	71	10	24
Soap / mini series	630	7	13	74	10	18
Competitive entertain	70	5	64	88	6	73
Movies	752	4	9	76	5	12
Sports	231	4	24	84	4	28
Average		14	29	80	17	35

Note. Viewer must be present for 30 seconds before the ad breaks. % is subject to rounding error.

χ^2 for *EOS / Presence* was calculated by using *EOS only* by *Presence* minus *EOS only*

χ^2 for *Total EOS / Presence* was calculated by using *Total EOS* by *Presence* minus *Total EOS*

Table 11 also shows the proportion of time viewers were oriented to the screen during their time present in the room. Overall there were significant differences between all the programmes types in terms of *EOS only:Presence* ($\chi^2 = 111.4255$, $df = 11$, $p < 0.01$) and *Total EOS:Presence* ($\chi^2 = 145.1223$, $df = 11$, $p < 0.01$).

Table 10 demonstrates that programme types differed in their measures of *EOS only* and *EOS + behaviour*. Table 11 compares *EOS only* with *Total EOS* (i.e.

any visual attention to the screen). It has been shown that the range of variation in the scores for *Total EOS* was also large across the programme types (9% to 64%). This means there are clear distinctions between some programmes. Using EOS as a indicator of viewers' involvement, programmes that have high *Total EOS* scores are therefore better able to hold the attention of viewers than programmes with lower *Total EOS* scores. An inspection of the results shows the competitive entertainment genre has the highest *Total EOS* score (64%) and is therefore best able to hold the attention of ad viewers. In contrast, movies and sports (4%) had the lowest *Total EOS* scores and one might conclude these genres are least able to hold the attention of ad viewers.

Another notable finding is that the two measures of attention do not vary in a similar way. While some programmes (e.g. chat shows) have relatively high results for both measures (*EOS only* and *Total EOS*), others (e.g. game shows) have high *EOS only* but low *Total EOS*, some have low *EOS only* and high *Total EOS* (e.g. informative interest programmes), yet others have low *EOS only* and low *Total EOS* (e.g. movies). While viewers paid highest *Total EOS* to competitive entertainment type programmes (64%), they actually looked at the screen fully (*EOS only*) 5% of the time. In contrast, some programmes had both high *EOS only* and high *Total EOS* scores. Chat shows had *EOS only* and *Total EOS* scores of 35% and 43% respectively. Movies, on the other hand, had the lowest *EOS only* score (4%) and the lowest *Total EOS* score (9%). These results suggest viewers pay varying degrees of attention to different ads in different programme types. However, because the relationship between the proposed measures and receptivity to ad messages is unknown, the measures must be considered separately in relation to different programme type.

A similar finding is evident when viewing behaviours are considered in the context of presence (see the last two columns in Table 11). This is shown by two sets of ratios; *EOS only*:Presence and *Total EOS*:Presence. The first ratio represents the proportion of time viewers were looking at the screen during the

time they were present during the ad break, and not engaged in any other activity. The second ratio shows the proportion of total time viewers had their eyes on screen, whether they were looking only at the screen or looking and doing something else simultaneously.

The range in the *EOS only* scores (4% to 35%) was much greater than the range in the presence scores (70% to 90%). If full "looking" (*EOS only*) was used as the criterion for programme selection, programmes would rank differently from when "presence" was used as the measurement. Even more important, the conclusion one might reach about viewing behaviour is that it is affected, not only by also by whether one uses *EOS only* or *Total EOS*, but whether these are considered as a proportion of the total time the ads were playing (i.e. includes non-presence in the room) as in Table 10 or considered as a percentage of time present. For example, the programme types with the first three highest *EOS only* scores were chat shows (35%), comedies (21%) and game shows (20%) (see column 2 in Table 11). But the three programme types with the highest percentage of viewers "present" were informative interest (90%), competitive entertainment (88%) and drama (87%) (see column 4 in Table 11). This demonstrates how programme types can be ranked in different ways depending on which measure is used. Thus measures of attention (i.e. *EOS only* or *Total EOS*) and measures of presence are not equivalent.

Programmes high in presence did not necessarily have high attention scores, whether attention was measured as *EOS only*, or as *Total EOS* (any visual attention). This is further illustrated by the data from Table 11. For example, during informative interest programmes, viewers were present for about 90% of the time (the highest presence score) yet only around 12% of viewers were actually paying full attention to the screen (*EOS only*). In contrast, chat shows tended to have the lowest number of people "present" but had higher *EOS only* than most other programme types. Viewers present during chat shows, therefore, tended to pay "fuller attention" to the screen than viewers of most

other programme types, despite having a lower proportion of audience present than most other programmes. Yet other programmes had high attention and high presence scores (e.g. chat shows and comedies). It is clear presence does not necessarily mean attention, and their relationship varied across different programme types.

The relationship between attention and presence is summarised by the ratios of *EOS only:Presence* and *Total EOS:Presence*. In considering this relationship, programme types varied differently again. For example, across the programmes, viewers were most attentive (measured by *EOS only:Presence*) during ads in chat shows (50%) and least during ads in competitive entertainment programmes and movies (5%). But when considering *Total EOS:Presence*, viewers spent the largest proportion of time looking at ads in competitive entertainment (73%) programmes and the least proportion of time on ads in movies (12%). These variations occurred because viewers were involved in different activities during different programmes and paid varying degrees of attention to the screen. The differences in the amount of attention paid to the screen showed the quality of the viewing and could be used to supplement information supplied by presence data. If presence data were used as a single variate measurement for "viewing", this information would be missed.

The measure of "presence", therefore, may not be a suitable surrogate measure for attention unless programme type is taken into account. If presence was a good indicator of attention, there would be a high correlation between attention and presence across the programme types. This was not the case.

This lack of consistency in the relationship suggests the two measures should always be considered separately in the context of the programme types. If there was a constant relationship between presence and attention for all the

programme types, it would be possible to weight presence data (or those obtained via ratings data) for each programme type. This was not the case.

There is also the possibility that even within each programme type, the relationship between attention and presence may vary across individual programmes. Just as viewing behaviour changes during different programme types, the quality of viewing may vary from programme to programme, even in the same genre category. To test out this possibility the ad-viewing patterns displayed by the audience during the programme category "comedy" were analysed. Table 12 demonstrates the four types of viewing behaviour of a range of comedy programmes. As expected the results show the ad viewers reacted differently in different ad situations. The same conclusions made about different genres can also be made about these individual programmes. The variation in terms of attention paid to the screen (*EOS only*), being involved in simultaneous activities (*EOS + behaviour*) and presence was as inconsistent as that found in the results for different genres (as in Table 10). These inconsistencies suggest that it may not be appropriate to assume that all programmes within the same programme type category make similar impacts on the ad viewer.

Table 12: Viewing behaviour during comedy programmes

Comedy programmes	No. of ads observed	<i>Present</i>			<i>Not present</i>	Total %
		%		%		
		<i>EOS</i> <i>only</i>	<i>EOS +</i> <i>behaviour</i>		<i>No EOS</i>	
		%	%	%		
South Park	26	46	6	48	0	100
Fresh Prince of Bel Aire	14	43	2	33	22	100
Friends	63	38	15	44	3	100
3rd Rock	31	34	23	30	13	100
Drew Carey Show	31	32	1	51	16	100
Sabrina	27	29	43	14	14	100
Will & Grace	32	21	18	36	25	100
The Simpsons	105	19	9	62	10	100
Pio	16	16	10	56	18	100
Newsflash	31	15	45	7	33	100
Dharma & Greg	57	12	7	53	28	100
Chalk	17	9	3	88	0	100
Wings	21	9	1	87	3	100
Keeping Up Appearances	8	7	4	89	0	100
Men Behaving Badly	18	7	3	83	7	100
Only Fools and Horses	8	7	29	64	0	100
Party of Five	16	7	1	5	87	100
Average		21	13	50	16	

4.6.2 Summary

There were significant differences between the eleven programme genres in viewing behaviours ($\chi^2 = 366.28112$, $df = 33$, $p < 0.01$). These differences were also observed for the ratio of attention:presence ($\chi^2 = 145.1223$, $df = 11$, $p < 0.01$). Programme types differed widely in the type of attention paid to ads in them. The rank order of the programme types also changed depending on the visual attention criteria used. One programme type might be high on *EOS only* and low in *Total EOS*, another might be high in *EOS only* and high in *Total EOS* suggesting that each programme type should be considered separately in relation to each attention criterion.

However, regardless of which visual attention measure (i.e. *EOS only*, *EOS + behaviour* or *Total EOS*) was used, the spread of the variation was rather wide. This suggests programme types could be separated by the amount of attention viewers paid to advertisements in each programme type. For instance, using *Total EOS* as a criterion, ads during programmes such as competitive entertainment programmes were better at "attracting" attention than others because they had the highest *Total EOS* scores. Advertisements during Movies, on the other hand, were the least able to attract attention.

When the measures of attention (*EOS only* and *Total EOS*) and the measure of presence were analysed together, there was also a wide variation between programme types. In particular, some programmes, which had high presence scores, did not necessarily have high attention (both *EOS only* and *Total EOS*) scores. An increase in presence in the room did not necessarily lead to an increase in attention to the screen. Furthermore, the relationship between attention and presence was not constant, as these measures varied in different ways for different types of programme.

As the relationship between attention and presence varied across different programme types, presence is not necessarily a suitable surrogate measure for attention or involvement, unless programme type is taken into account. This has implications for providers and users of ratings data (Peplemeter) since presence data are often used to represent "viewing" (attention) or viewers' commitment to television programmes.

However, the proposed solution of weighting measures of presence to account for differences in ad-viewing behaviour seems suspect, since individual programmes within a genre seem to be subject to considerable variation, as evidenced by Table 12.

4.7 Viewing Behaviour by Impact Categories Developed by Hoffman and Batra (1991)

This section examines viewer behaviour associated with programme types defined in terms of programme impact scores. These scores were derived from viewer ratings on the scales developed by Hoffman and Batra (1991). According to Hoffman and Batra (1991), programmes could be classified by their affective and cognitive effects on viewers. Another dimension was whether the impact was "high" or "low". On the basis that programme involvement was multi-dimensional, Hoffman and Batra (1991) proposed programmes could be grouped into three categories, "High Affective", "High Cognitive" or "Low Impact".

In this study, a new category, High Impact was added to the three impact categories to include a programme category where viewers were involved both cognitively and affectively (see p 77-79 for full explanation). These four categories—High Impact, High Affective, High Cognitive and Low Impact—were used in this study. The four impact categories were derived from whether viewers agreed that they had "learned" (cognitively) from the programme and whether their feelings were "touched" (affectively) by the programmes.

The core dimensions of these categories are summarised below:

- High impact programmes affect viewers both cognitively and affectively
- High cognitive programmes affect viewers cognitively but not affectively
- High affective programmes affect viewers affectively but not cognitively
- Low Impact programmes did not affect viewers cognitively or affectively.

The sections that follow report results of viewing behaviour by these four categories of programmes. Specifically, an examination is made of whether there were clear differences in ad-viewing behaviour for each programme type, and the value of classifying programmes in this way.

4.7.1 Viewing behaviour associated with programme impact scores

There were significant differences in ad-viewing behaviour between the four categories of programmes ($\chi^2 = 23.5963$, $df = 9$, $p < 0.01$) (see Table 13). This suggests that programmes with different impacts result in different viewing behaviour.

Table 13: Viewing behaviour by impact scores

Programme impact types	No of ads observed	<i>Present</i>			<i>Not present</i>	Total %
		<i>EOS only</i>	<i>EOS + behaviour</i>	<i>No EOS</i>	%	
High Impact	775	10	15	44	31	100
High Affective	566	8	6	71	15	100
High Cognitive	1213	16	13	45	26	100
Low Impact	2819	14	15	55	16	100
Average		12	12	53	23	100

Note: Viewer must be present 30 seconds before the ad breaks. % is subject to rounding error.

Three of the programme categories, High Cognitive, Low Impact and High Impact, had almost similar *EOS + behaviour* scores. Ad viewers were "fully looking at the screen" for about 16% of their time. Viewers in this category could be considered as "most attentive". For another 13% of their time, they were engaged in simultaneous activities, as they looked at the screen (measured by *EOS + behaviour*). In contrast, High Affective programmes had the lowest *EOS only* and *EOS + behaviour* scores (8% and 6% respectively). Ad viewers of this programme type hardly looked at the screen (*No EOS* = 71%). They were the least "attentive". But the results of the High Cognitive impact programmes were not very different from Low Impact and High Impact programmes. For example, the scores for *EOS only* and *EOS + behaviour* of Low Impact programmes were 14% and 15% respectively. These scores were different from those of the High Cognitive impact programmes by no more than 2%. High Impact programmes had a slightly lower *EOS only* score (10%) but the *EOS + behaviour* score (15%) was not very different from High Cognitive impact programmes (13%). Thus, programmes in the High Cognitive impact category were better at "capturing" the attention of ad viewers than High Affective programmes and actually indicated a very similar pattern of viewing behaviour to both Low Impact and High Impact categories. That programme types of high and low impacts involved ad viewers in similar ways, suggests the categories are not particularly useful for separating their viewing behaviour.

Moreover, contrary to expectations, programmes classified as High Impact (high in cognitive and affective impact) did not obtain the highest ad-attention scores. Viewers looked fully at the screen 10% of the time (*EOS only*) and looked and did something else 15% (*EOS + behaviour*) of the time. This shows that attention was low suggesting that only some parts of these programmes may have "attracted" some attention from their audience.

4.7.2 Visual attention during presence

So far the results have shown variations in the proportion of attention paid to the screen during the ad breaks for each programme impact type. Two other variables in the analysis were the proportion of time viewers were "out of the room" (not present) during the ad breaks and the total amount of attention paid to the screen during the ads. It was also possible to find the quality of the viewing (using ratios of attention:presence) for each programme impact category.

As shown in Table 14, the range of scores for the four categories in terms of presence varied by 16% (69% to 85%). Each of the categories gave rise to different expectations about viewing behaviour. For example, one would expect High Impact programmes to be the most "involving". Thus, if presence was used as a surrogate measure, High Impact programmes would have the highest "present" score and Low impact programmes the lowest score. These results, however, show this was not the case: the highest "presence" score was associated with the High Affective programmes (85%); the next highest was for Low impact programmes (84%). The difference was only 1%. High Impact programmes, in contrast, had the lowest proportion of viewers present (69%). The expectations for the different categories are not confirmed by the observations.

Table 14: Viewing behaviour by impact scores during time present

Programme impact types	No of ads observed	<i>EOS</i> <i>only</i> %	<i>Presence</i>			
			<i>Total EOS</i> (<i>EOS only</i> + <i>EOS +</i> <i>behaviour</i>) %	<i>(EOS only +</i> <i>EOS +</i> <i>behaviour +</i> <i>No EOS</i>) %	<i>EOS only</i> : Presence %	<i>Total</i> <i>EOS</i> : Presence %
High Impact	775	10	25	69	15	36
High Affective	566	8	14	85	9	17
High Cognitive	1213	16	29	74	22	39
Low impact	2819	14	29	84	17	35
Average		12	24	77	16	31

Note: Viewer must be present for 30 seconds before the ad breaks. % is subject to rounding error.

χ^2 for *EOS* / Presence was calculated by using *EOS only* by presence minus *EOS only*

χ^2 for *Total EOS* / Presence was calculated by using *Total EOS* by presence minus Total EOS

The relationship between the measures of attention and presence for the four categories is also interesting. The difference across the four programme types using *EOS only*:Presence was not significant ($\chi^2 = 4.6889$, $df = 3$, $p > 0.05$). Nonetheless, as seen from Table 14, on average, across the four impact classifications, viewers full attention (as measured by the ratio of *EOS only*:Presence) was a low 16%. Viewers in High Cognitive programmes tended to be most attentive (22%), while viewers of High Affective programmes were least attentive (9%).

By comparing this attention measure to presence, the results suggest that for some programme types, presence again did not guarantee attention. For instance during High Affective programmes, viewers were in the room for 85% of the time, yet *EOS only* :Presence was only around 9%. By contrast, viewers during High Cognitive programmes were present for 74% of the time yet were

more attentive during that time (22%). Viewers of High Impact programmes were present for 69% yet were only moderately attentive (15%). These results highlight that presence, as a measurement on its own does not provide a reliable measure of ad viewers' behaviour.

In comparing the ratio of *Total EOS:Presence*, the same trend was observed. Here the difference across the programme type categories for *Total EOS:Presence* was significant ($\chi^2 = 12.0592$, $df = 3$, $p < 0.01$). Specifically, this means the impact categories could be differentiated by the aggregate of total time spent looking at the screen. As indicated in Table 14, using *Total EOS:Presence*, High Affective programmes seemed very different from the other three types. This programme type had viewers who paid the lowest amount of attention to the set (17%) and differed widely from the other three; programme scores were twice as high as that for High Affective. The scores for the other three programme types were relatively close (35% for Low Impact, 36% for High Impact and 39% for High Cognitive).

It is apparent from the results that High Affective type programmes are different from the other three categories, all of which have similar *EOS only:Presence* and *Total EOS:Presence* scores. There seemed only weak support for differentiating the programme categories by impact. The next section examines this relationship further by comparing the results of the present study with those obtained by Hoffman and Batra (1991).

4.7.3 Comparing results on impact categories

The present study replicated the labels used in the classifications by Hoffman and Batra (1991) and added a fourth, High Impact, category. As this was a new category, there was no direct comparison for this impact type and it was not possible to derive direct expectations for these High Impact programmes. A

slightly modified method was also used here to derive the results for the four categories. As slightly different variables and measurements were used, the expected results may be different from those of Hoffman and Batra (1991).

Tables 15 and 16 provide a comparison between the findings by Hoffman and Batra (1991) and those of the present study. Table 15 shows the summarised results of attention and presence in Hoffman and Batra (1991). No specific percentage for the measures was provided in that study. Only rough indications such as "all attention", "some attention" and "hardly any attention" were given. Table 16 shows results of the percentages of attention and presence for each programme type in the present study. It also provides a rough indication of how programme types rank in relation to each measure. The two tables only appear to have direct comparative references at face value and are presented here to show broad differences. Moreover, only the first three classifications of Table 15 could be used to make comparisons with the viewing behaviour shown in Table 16.

Table 15: Viewing behaviour by Hoffman and Batra (1991)

Programme impact type	Viewing behaviour Hoffman's and Batra (1991)	
	Attention	Presence
High cognitive	All attention	High—do not leave room
High affective	Some attention	Low—Leave room during ads
Low impact	Some to hardly any attention	Low—More likely to leave room

Table 16: Viewing behaviour in the present study

Programme impact type	Viewing behaviour (this study)			
	Attention			Presence
	<i>EOS only</i>	EOS + behaviour	Total EOS	
High cognitive	Highest 16%	Moderate high 13%	Highest 29%	Moderate high 74%
High affective	Lowest 8%	Lowest 6%	Lowest 14%	Highest 85%
Low impact	Moderate high 14%	Highest 15%	Highest 29%	High 84%
High impact	Moderate high 10%	Highest 15%	Moderate high 25%	Lowest 69%

Hoffman and Batra (1991) claimed that High Cognitive and High Affective programmes involved viewers highly and their viewers reacted to these programmes differently in terms of attention and presence. There were also differences in behaviour during the programme versus during the ad breaks. Nonetheless, although the authors did not measure the proportion of time viewers spent on different activities during the ad break, they generalised that during High Cognitive programmes, viewers tended to give full attention to the programmes, did not talk and did not leave. This was supported in the present study as viewers paid the highest proportion of attention measured by *EOS only* (16%) and *Total EOS* (29%) to ads in these programmes too. Viewers were also present during the ad break for a moderately high proportion of the time (74%).

Hoffman and Batra (1991) also suggested advertising would be most effective during High Cognitive impact programmes because there would be more opportunities for viewers to be visually "involved" and to be thinking about the ads during these programmes. Again, there was support for this in the present study as High Cognitive programmes had the highest proportion of *Total EOS:Presence* (39%) as well as *EOS only:Presence* (17%).

Hoffman and Batra (1991) regarded viewers of High Affective programmes as not very attentive : viewers would give some attention to these programmes but during the ad break, presence was low. In the present study, however, while attention to ads in such programmes was lowest among the four programme types—*EOS only* (8%) and *Total EOS* (14%)—they had the highest presence score (85%). Therefore, it was necessary to examine the results of High Affective programmes further by looking at their attention:presence ratio.

During High Affective programmes, viewers tended to display the lowest visual attention when they were in the room. Both *Total EOS:Presence* (22%) and *EOS only:Presence* (9%) were lowest among the four programme types. Hoffman and Batra (1991) concluded that the probability of the viewer staying in the room and watching the ad was not very high. In the present study, although High Affective programmes had the highest presence score (85%), of the viewers who stayed in the room, they were not often attentive to the screen.

For Low Impact programmes, the authors argued that viewing behaviour was more or less "random" and viewers were therefore unlikely to process closely either the programme or the advertisement suggesting that advertisements placed during such programmes might not be seen or would be poorly "received". In this study, in contrast, Low Impact programmes had relatively higher *EOS only:Presence* (17%) and *Total EOS:Presence* (35%) compared with the other three—meaning attentive levels were higher during these programmes. Using attention as a measure, there was little support for the claim

that advertising would be less effective during Low Impact programmes, as suggested by Hoffman and Batra (1991).

4.7.4 Relationship between programme types by genres and impact categories

The previous sections showed that ad viewing differed from programme to programme and arrived at different conclusions from Hoffman and Batra (1991). However, in both studies, respondents were required to state, via a programmed impact rating scale, how the programmes had affected them. This method posed many problems. One major problem was viewers perceived programmes as having different levels of impact, making it difficult to categorise them by this factor, which raised questions of validity. To show these complications, Table 17 presents details of how viewers rated programme genres by each of the four impact categories.

For some programmes, the relationship between genres (programme types) and impact categories was straightforward and obvious. Programmes such as comedies, soaps, chat shows and movies could be grouped as Low Impact type (see Table 17). They all have percentage ratings as a Low Impact programme of above 80%. In other words, most viewers felt that these programmes did not impact on them cognitively or affectively even when between 7% to 16% gave impact ratings which would have classified the genre as High Affective. For some other programmes, however, the distinction was less clear. Sports programmes, for example, could be termed as both High Affective (41%) or as Low Impact (46%) as they had almost equivalent scores for both. And for 13% the sports programmes were classified as High Cognitive on the basis of the viewer ratings.

Table 17: Impact scores by programme types

Programme type	No. of raters	Impact classifications (Mean %)				Total %
		High Impact %	High Affective %	High Cognitive %	Low Impact %	
Soap	26		7		93	100
Chat	5		15		85	100
Movie	23	5	7	1	87	100
Comedy	43	3	16	1	81	100
Sports	19		41	13	46	100
Drama	54	13	23	1	63	100
News	116	31	1	52	17	100
Informative Interest	17	33	7	26	34	100
Musical	13		12	25	64	100
Game	12		32	40	28	100
Competition entertain	6			72	28	100

% is subject to rounding error

Other programmes were even more widely spread across the categories. Drama was classified as either Low Impact (63%), High Affective (24%) or High Impact (13%) by relatively high proportions of viewers. News and informative interest programmes were equally problematic, with ratings of High Impact (31%), High Cognitive (52%) and Low Impact (17%). Many programmes could not be classified as belonging only to one fixed category thus calling into question the validity of the classification scheme.

4.7.5 Summary

There was some statistical evidence that the four impact categories differed significantly across the four viewing behaviours ($\chi^2 = 23.5963$, $df = 9$, $p < 0.01$) and for *Total EOS:Presence*, ($\chi^2 = 12.0592$, $df = 3$, $p < 0.05$), but not for *EOS only:Presence* ($\chi^2 = 4.6889$, $df = 3$, $p > 0.05$). An inspection of these distinctions, however, suggested that three of the four categories provided very similar results.

Low Impact programmes (such as those that did not impact either cognitively or affectively) were rather similar to High Cognitive programmes in terms of viewing behaviour and also to High Impact programmes. The scores for *EOS only* varied slightly (by 2%), while for *Total EOS*, they were equivalent (both at 29%). Even in terms of the ratios of attention to presence, the variation was small. High Impact programmes, on the other hand, had moderately high attention scores suggesting some parts of these programmes attracted more attention than other parts. Low Impact programmes were able to sustain high attention and were similar to programmes that cognitively impact on viewers.

While High Affective programmes had the highest presence scores, they had the lowest *EOS only:Presence* and *Total EOS:Presence* scores—there was an inverse relation between attention and presence. Presence did not guarantee attention, and in this case, impact scores were not related to the measure of presence.

There were both similarities and differences in the scores for viewing behaviour between this study and that of Hoffman and Batra (1991). As in Hoffman and Batra (1991), while viewers paid higher attention and were present longer for advertisements in High Cognitive type programmes, they paid lower attention to advertisements in High Affective programmes. Contrary to their findings, however, presence was highest during advertisement breaks in High Affective

programmes. Furthermore, the present study found viewers to be highly involved in advertisement breaks during Low Impact programmes rather than least involved, as suggested by Hoffman and Batra (1991). In terms of viewing behaviour, therefore, this study contends that advertising would work better in High Cognitive programmes than in High Affective and Low Impact programmes.

It should also be noted that programme genres did not fit neatly into clear-cut impact categories. Some programme genres fell into more than one impact category, sometimes even into opposite impact types and intensity. The implication is that either individual programmes varied within genres, or individual viewers' perceptions are quite different for the same programmes—an implication that could make it difficult, if not impossible, to match individual programmes and genres to impact categories.

This chapter has presented the major results of the study and pointed out the validity of using attention measures (eyes-on-screen behaviour) to show the level of viewer involvement and to provide valuable information about what actually happened during the ad break. The next chapter discusses the main issues raised by the findings. It points out the implications of the results for media practitioners and also discusses the limitations of the study.

CHAPTER FIVE

DISCUSSION

5.1 Introduction

Advertisers invest large sums on television advertising in the hope that the masses of viewers who watch television will be positively affected by their advertisements. Advertisers commonly use the "ratings" of a TV programme to indicate the size of the viewing population "present" and therefore "exposed" to their ads. But it has been suggested that the information provided by ratings is misleading, inaccurate and superficial.

The advertiser's interest is obviously in the effectiveness of a particular advertising spot. However, even when ad ratings are provided for an advertising spot, they indicate only the size of the audience present during the screening - but fail to distinguish "attentive" ad viewers from non-viewers. Rather than being a dichotomous distinction of attentive and non-viewers, during an ad break viewers may engage in many different activities, all of which can be performed while the viewers are in the room and recorded as being present. This suggests that there are serious methodological problems in using presence as a measurement of advertising effectiveness.

In addition, the effect of different programmes on ad-viewing behaviours has been the subject of speculation. The problems introduced by a potential bias in ratings data are severely compounded if different programmes have different effects on the ad-viewing behaviour.

However, if ad-viewing behaviour does not differ for different programmes, then the potential bias in estimating ad-viewing audience size is much less of a problem to advertisers. This is clearly the case for decisions about which programme to advertise in; a programme with twice the audience size compared with another programme would have twice the ad-viewing audience of the other programme and thus any media-planning decisions would be unbiased. For comparisons of different media, any overstatement of audience sizes for television viewing might mean that other media are disadvantaged in terms of media decisions. Thus, for example, if television audience sizes are overstated, then this may mean that print media will be disadvantaged in media buying decisions, presuming of course that the audience size for print media are not similarly overstated. An approach to this problem might simply be to discount the audience size estimates obtained from ratings data.

If different programmes do have different effects on ad-viewing behaviour then an obvious solution is to provide an adjustment to the size of the audience estimated from ratings data. This may be done qualitatively for each programme; in fact it might be necessary that this be done for each episode, or at a more minute level, each segment of the programme preceding the advertisement.

Alternatively, a more tractable solution might be to test whether this adjustment could be done at a more aggregate level, say for each programme type. It would obviously be easier for media planners if these adjustments could be done for whole programme categories. In fact, media planners might make these adjustments qualitatively already. In this case, this study would be valuable for examining whether these qualitative judgments are correctly calibrated.

Programmes can of course be grouped in a variety of ways. The present study investigated three programme classification systems used in earlier studies of television viewing behaviour. If a reliable relationship between the programme

classification and ad-viewing behaviour could be found, that is if programmes of a given type have small audiences, but much more favourable ad-viewing behaviour, then this might lead to a different media decision than if audience size alone was used.

The purpose of this study was to research viewing behaviour, specifically by examining viewing behaviour during advertisements. In addition, the study sought to explore the relationship between the programme that preceded the advertisement and ad-viewing behaviour.

In the present study, an approach of trying to gauge quantitative elements of audience behaviour during ad breaks is used. This was done by cataloguing at 3-second intervals, viewer behaviour during ad breaks into four possible states:

- Eyes on screen (*EOS only*)
- Eyes on screen and doing something simultaneously (*EOS + behaviour*)
- Eyes not on screen (*no EOS*)
- Not present

However, it must be recognised that such measurements are surrogates for the advertiser's ultimate interest in the effectiveness of a particular advertisement. While the advertiser's interest is obviously in comparing the effectiveness of two or more presentations of an advertisement, (i.e. is it better to advertise in Programme A rather than Programme B?), the measurement of advertising effectiveness is the subject of intense debate. Effectiveness could be measured in terms of a sales effect, or a myriad of surrogate indicators, such a change in attitudes, buying intentions, recall, recognition or in fact any of the measures of advertising effectiveness that have been proposed in the advertising literature.

The major findings in this research can be summarised as follows:

Visual attention to the screen was generally low, despite viewers being present for a high proportion of the time during the ad breaks. Viewers were present, on average, during the ad breaks for 78% of the time, but were only paying attention to the screen fully 14%, and looking as well as doing something else simultaneously 11% of the time. On average, 52% of the time, no viewers were looking at the screen.

Not only was the proportion of high quality ad viewing low, as represented by only a quarter of observations of ad viewing being coded as visual attention being directed to the screen, this proportion differed by programme type. For example, some programmes could command as much as ten times the proportion of eyes-on-screen than other programmes.

However in examining three specific methods of classifying programmes into types, none of them provided workable descriptions of ad-viewing behaviour. This problem manifests itself in a number of ways. Primarily, any substantive conclusions differ according to the way that the ad-viewing behaviours are summarised. In addition, aggregations of programmes into categories mask considerable within-group variation.

For example, programmes classified as *demanding* and *less-demanding* programmes did not differ significantly in ad-viewing behaviour. However, when considering the amount of visual attention given to the screen during the time viewers were present in the room, *demanding* and *less-demanding* programmes did differ significantly. Furthermore, the viewing behaviour in the *less-demanding* category varied across different programmes within that category in a statistically significant way. Some even had very similar patterns of visual attention to *demanding* programmes.

When classifying programmes by genres, again different conclusions are reached depending on which measure of ad-viewing behaviour is being considered. Genres ranked differently when attention measures were used rather than presence. Furthermore, genres that ranked highly in terms of presence may not contain very attentive ad viewers, in terms of the proportion of viewing classified as *EOS*.

Using impact categories to differentiate between programmes, viewers in High Cognitive impact programmes were most "attentive" to ads while viewers in High Affective impact programmes were "least attentive".

The sections that follow discuss these major findings explored in this thesis. Following this, the limitations, and the implications for future research, are identified and discussed.

5.2 Presence Versus Attention

When viewers turn on the TV, there is a high probability that they will also stay in the viewing room. In the 1960s, Ehrenberg and Twyman (1967) claimed viewers were present 80% of the time the TV was turned on. In the 1980s, using video recording research, Collett (1986) reported similar results (i.e. 80% presence when the TV was on). Another study by Svennevig (1987), also using video observation, reported viewers were present for 70% of the time the TV was on. Results in the present study found a similar level, although slightly higher level of presence, with a viewer being present 85% of the time the TV was on.

However, these studies have aggregated results for viewing of programmes and viewing of advertisements. Advocates of Peoplemeter rating systems have consistently claimed a high proportion of viewers watch ads. Using Peoplemeter data, Danaher (1995) claimed 90% of programme viewers were also present during ad breaks. Cook (1995) also reported viewers were present 95% of the time during commercials, but acknowledged that viewers were involved in many different activities such as talking, reading, playing, doing chores, and even sleeping. However, the methods used in obtaining these figures are subject to considerable criticism, on the grounds that they, at best use people's self reports of behaviour.

Results in the present study, albeit using a different methodology than that used in obtaining the typical Peoplemeter estimates, indicated the proportion of time viewers were present during the ad breaks, was on average, 78%. Thus, Peoplemeter derived estimates appear to claim higher estimates of the audience size for ad viewing than those obtained from this study's observations of actual ad-viewing behaviour.

One interpretation of these two separate estimates might, in fact be, that they are not that significantly different, although this is an arguable point in considering the efficiency of a multi-million dollar media plan. However, any interpretation of these estimates has to consider that the viewers who are recorded as present may, in reality, be doing things that accompany the viewing, or they may be involved in activities that completely take their attention away from the screen. If the measurement of presence alone is used, much of this information is overlooked.

The results from the present study indicate that ad viewing is often accompanied by other activities (as noted by Cook, 1995; Zwaga 1992a; Collett 1986; Svennevig, 1987). The viewer watching the screen with total undivided attention, totally attentive to the screen was not the norm in this study. Uninterrupted attention to the television screen during the commercial break was rare. Most people did something else during the majority of the time they were in the room during commercial breaks. On average, across all the ad breaks, *EOS only* was around 14%; for a further 11% of the time, viewers were simultaneously involved with other activities while looking at the screen (*EOS + behaviour*). This gives a maximum actual "looking" time of just 25%.

This is much lower than the findings of Collett (1986) who reported that viewers were actually "looking" at the screen 65% of the time. Svennevig (1987) reported viewers were partially attentive to the screen for around 20% of the time, and fully attentive 27% of the time, which means their total "looking" time was 47%—again much higher than the present study. However, Collett (1986) and Svennevig (1987), however, were reporting results on the attention paid to programmes and not specifically to ads.

Using a measure similar to the present study, that of EOS, Krugman, Cameron and White (1995) found viewers had their eyes-on-screen 33% of the time ads

were on. But of those visually attending, less than 50% had eyes-on-screen of more than 50% of the time.

The findings of the present study were also more consistent with another study conducted in New Zealand in 1992. Although not measuring time spent on ads, Colmar Brunton Research (1992) found that only 8% of ad viewers reported actually watching and listening to the advertisements. During the ad break, 47% of viewers had physically left the room, 30% were talking or doing something else, and 12% had switched channels.

These findings show attention to advertisements was generally not high. When viewers were looking at the screen, their visual attention was often divided equally between looking and being engaged in some other simultaneous activity. These incidences of "divided attention" may include reading a book and occasionally looking up to glance at the television, or watching part of the time and keeping an eye on the children, the rest of the time.

5.3 Viewing Behaviour and Programme Classification Systems

As previously mentioned, the findings discussed in the preceding section are probably familiar to most media planners; in fact few media planners would believe that their ad was actually being exposed to the entire audience suggested by ratings data.

A specific concern of this thesis was to investigate whether the proportion of time viewers spent visually attending to advertisements varies with the type of programme being screened. There may be certain types of programmes that appear to hold viewers' attention better than others, and, in the context of advertising there may be some residual effect, which carries over into the ad break. Alternatively, a programme, which does hold viewers' attention, might have a residual effect such that viewers take a cognitive break thus negatively impacting on ad-viewing behaviour. Such suggestions are obviously speculative, pointing to the need for empirical work to resolve such debates. Thus if ad-viewing behaviour differs by different programme types, then this presents a need for media-planning decisions to be made in light of this observation

Three programme classification systems were examined in the present research. The first programme classification method was developed by Barwise and Ehrenberg (1988), the second by Danaher (1995) and the third by Hoffman and Batra (1991). When the present study compared these different programme classification methods in terms of the behaviour of ad viewers, it was found that attention to advertisements varied for different programmes in a way that was not consistent with the ways programmes were classified. Thus, while the classification methods may have been based on consistent relationships with certain aspects of viewing behaviour, this study could not find consistent relationships between these classifications and ad-viewing behaviour.

The sections that follow discuss the implications of the results derived from each approach and conclude by highlighting the main issues common to all three.

5.3.1 Barwise and Ehrenberg (1988) classification

Barwise, Ehrenberg and Goodhardt (1979) and Barwise and Ehrenberg (1988) found consistent results in the appreciation scores viewers gave to programmes for many standard programme types, and decided to group them into two larger groups: Information and Entertainment. Information programmes were considered to be *demanding*. It was felt viewers put more effort into watching them than into Entertainment programmes, which viewers watched for relaxation. Entertainment programmes were considered to be *less-demanding*.

The present study investigated how ad viewing related to the programme types categorised in this way. Thus the intention was to test whether viewers who were more involved in *demanding* programmes paid more attention to the ads in them.

The present study found no significant differences between the categories of *demanding* and *less-demanding* programmes when comparisons were made using the four viewing behaviours (*EOS only*, *EOS + behaviour*, *No EOS and not present*). However, *demanding* programmes did differ significantly with regard to the amount of time viewers spent attending to the screen as a proportion of time present during the ad break from *less-demanding* programmes. In other words, the two categories could be differentiated by the index *Total EOS:Presence* but not by the four ad-viewing behaviours.

However, the observation that there is no difference between the two categories (*demanding* and *less-demanding*) in terms of the pattern of the four viewing

behaviours, was not observed for programmes within the *less-demanding* category. In addition, although there was considerable variation in the *less-demanding* programmes, this pattern was not evident for the programmes in the *demanding* category.

The two categories of programmes within *demanding* programmes; news and documentary differed little in the patterns of ad-viewing behaviour. In contrast, across the four *less-demanding* programmes there were significant differences in the patterns of viewing. For example, light entertainment and light drama programmes had *EOS only*:Presence of 19% and 17% respectively. These scores were very much higher than those of films and sports (each at 5%). Thus, it might be concluded that light entertainment and light drama were accorded patterns of ad-viewing behaviour much more similar to *demanding* programmes than to films and sports.

However, such a conclusion differs according to the measure of ad-viewing behaviour used. This could be because of variations in content, which made varying impacts on the viewers. This was not surprising, as according to Barwise and Ehrenberg (1988), some *less-demanding* programmes have characteristics that were more *demanding* than those programmes, which might be more properly identified as *demanding* ones.

Moreover, Barwise and Ehrenberg (1988) also stated that different entertainment programmes have different purposes and affected viewers differently. While some programmes may only entertain, others both entertain and inform. In Barwise and Ehrenberg's (1988) study, the level of demand placed on each programme was subjectively determined (by the viewer) and depended on the viewer's perception of the demands put to them. In the present study this was also evidenced by the variations in the degrees of attention paid to the four *less-demanding* programmes.

In summary the classifications of programmes into *demanding* and *less demanding* (as defined by Barwise and Ehrenberg (1988)) do not seem appropriate in describing the attention viewers pay to ads. The implication for television advertisers is the need to look beyond these two (rather broad) categorisations of programmes and assess how individual programmes make demands on the viewer.

5.3.2 Genre classification

The previous section discussed how researchers grouped programmes by the presumed level of effort demanded of viewers in watching programmes. An alternative approach is to expand the classification from two categories and to examine how ad-viewing behaviour differs by genre. This section discusses the findings on the viewing behaviour for programmes classified in this way.

In analysing ad-viewing behaviour across different genres, results showed there was a large variation, and this difference was statistically significant. Thus it can be concluded that viewers paid varying degrees of attention (*EOS only* or *EOS + behaviour* or *Total EOS*) to different genres and the difference in the computed attention indices was wide.

However, again different conclusions would be reached as to which genre would be *best* to advertise in, depending on the measure of ad-viewing behaviour used. Although on average, a genre may have high presence percentage, viewers may not have paid high visual attention to the screen. For example, informative interest type programmes have a high number of ad viewers present (90% of the time) but viewers only paid 12% attention (*EOS only:Presence*) to the ads in them. One might conclude that the competitive entertainment programme is the best genre to capture the attention of viewers as it has the highest score for *Total EOS:Presence* among the programme types.

only, this highest score was claimed by chat shows. In other words, when separating programmes by genre, it is important to consider which attention measure to use.

These conclusions have to be contrasted with those reached, for example, by Danaher and Lawrie (1998) who reported that viewers' commitment scores to programmes were as high as 78% (measured by PMV) and 60% (measured by P80+). PMV is the proportion of advertising minutes viewed of the total ad time viewers could have watched. P80+ is the criterion that the viewer must be present for at least 80% of the time the ads were on.

The present study found ad-attention rates to be comparatively low measured by *Total EOS:Presence* (30%) and *EOS only:Presence* (15%). In fact the attention scores (measured by *Total EOS:Presence*) for four programme genres—(news, soap, game show and drama)—were only half as high as the commitment scores of Danaher and Lawrie (1998) for the same four programme genres. Danaher and Lawrie (1998) also suggested that information type programmes such as news showed the most committed viewers. By contrast, using attention measures (by both *Total EOS:Presence* and *EOS only:Presence*), the present study found entertainment type programmes such as chat shows and game shows had the most attentive ad audiences.

The relationship between presence and attention may also not be constant within each genre category. Although not explicitly tested for all genres, it was found there were large variations in viewing behaviour across different programmes within the comedy genre. It is possible that similar results could be obtained when other genre categories are analysed. In other words, the influence of individual programmes on viewers should not be discounted. Within a group of similar type programmes, different impacts are made by the individual programmes and the ad-viewing behaviour is quite different.

5.3.3 Hoffman and Batra (1991) classification

In the previous sections, programmes were grouped by programme genres, namely by using subjective judgments to assign a particular programme to a genre, and in the case of Barwise and Ehrenberg's (1988) method, further assigning genres of programmes to two hierarchical categories; *demanding* and *less-demanding*. In contrast, the programme classification system developed by Hoffman and Batra (1991) relies on viewers' subjective ratings of the programmes.

Three impact categories were identified by Hoffman and Batra (1991) in terms of whether the programme was classified as low intensity or whether the intensity was high-affective or high-cognitive. In the present study, it was not possible to utilise these categories because a different methodology was used to classify participants' responses. Instead of using their three categories (High Affective, High Cognitive and Low Impact), four similar categories (High Impact, High Affective, High Cognitive and Low Impact) were developed.

Programmes that viewers "learned from" and "touched their feelings" were classified as being High Impact programmes. Those programmes that viewers "learned from" were classified as High Cognitive programmes, while those programmes that "touched (their) feelings" were classified as High Affective programmes. The programmes that neither "touched (viewers) feelings" nor those they "learned from" were classified as Low Impact programmes.

The results of the present study demonstrated significant differences in the viewing behaviour within the four impact categories. Viewers in High Cognitive impact programmes were most "attentive" to ads, while viewers in High Affective impact programmes were "less attentive". During the ad break, viewers in High Cognitive and Low Impact programmes spent almost one-third of the time (29%) "looking" at the screen. In comparison, in High Affective

programmes, only slightly more than one tenth (14%) of that time was spent looking at the screen.

The results of the present study support Hoffman and Batra's (1991) results for High Cognitive programmes but found differences from their results for High Affective programmes. In the present study, viewers were present during High Cognitive programmes for about three-quarters of the time (74%) and spent a high proportion of the time attending to the screen. Using total visual attention over presence (*Total EOS:Presence*), ad viewing during High Cognitive programmes (39%) was also much higher than the other three categories. These results support Hoffman and Batra's (1991) description that viewers of High impact programmes tended to give all their attention to the programmes and ads during viewing.

During High Affective programmes, in contrast, viewers in the present study paid the least visual attention but were less inclined to leave the room during the ad break (*not present* = 15%). When using presence to represent "viewing", High Affective programmes had the highest rating score. Again, this is a good example of a case when presence does not equate with attention. Although viewers in High Affective impact programmes were present for the ads for around 75% of the time, they paid little attention to the screen (17% of the time). The activities the viewers were engaged in were perhaps those that Hoffman and Batra (1991) described as "competing" with the viewing experience: where attention was low and there were other distractive behaviours during the viewing. This suggests that perhaps advertisers would be short-changed if ratings alone were used as the main measure for advertising effectiveness during High Affective programme types.

Hoffman and Batra (1991) also suggested that their high impact programmes are high in intensity of feelings and cognitions. For example, viewers' intensity could be so high during High Affective programmes; they might leave the room

during the ads to release their pent-up "emotions". This however was not shown in the present study. Viewers of High Affective programmes tended to stay on during the ads.

Contrary to Hoffman and Batra's (1991) claims that Low Impact programmes were least effective for advertising, in the present study, programmes evaluated as Low Impact secured around 30% of total visual attention to the screen (*Total EOS*). While viewers of this type of programme were not present in the room for 84% of the time, those who did stay spent 35% of that time looking at the ads. The implication is that even Low Impact programmes were able to "hold" the attention of the viewers for at least one-third of the time they were present. Therefore, despite those programmes classified as having the least impact, ads screened during these programmes were being still "watched" often. In fact, using *Total EOS:Presence*, the score for Low Impact programmes (35%) was as high as that of High Impact programmes (36%). Therefore ad-viewing behaviour does not seem to correlate with the expectations suggested by the impact categories.

While these are promising early results, the measurement of impact-scores introduces considerable problems in implementation, since it would require separate viewer ratings of impact for each programme. If there were a high correspondence between impact categories and genres, it would be possible to use genres to differentiate between the four types of impacts. But the present study did not find supporting evidence for this. Some genres were found to belong to more than one impact category, which negates the purpose and undermines the value of classifying programmes into larger groupings.

Each programme should rather be treated by its own impact score or involvement criterion. Instead of grouping programmes according to fixed impact categories, programmes should be rated separately, based on their own merit. One way to do this is to conduct random (next-day) evaluation of each

separate programme screened using a large sample base of viewers. In this way, advertisers would be able to know the direct value each programme has for the viewers. Once that is known, advertisers could then select combinations of programmes most likely to be viewed by their targeted market segment or product users. Media buyers would be able to select individual programmes (or groups) based on the type of impact and response they wanted from their viewers, and advertisers would feel confident about their buying choice and be able to justify their media decisions.

5.4 Conclusion and Contribution

Overall, the results in this study demonstrate that while presence was high (78%) during the ad breaks, average visual attention to the advertisements on television in most programmes was low (never more than 25%). During the advertisement breaks, viewers were often engaged in many activities other than only looking at the screen. However, since no media-planner seriously believes that the audience size is exactly the number indicated by ratings data, the contribution of this thesis is in providing a basis for calibrating media planners' own subjective and qualitative judgments regarding this issue.

While the overstatement of audience sizes need not be alarming to those who make advertising decisions the indication that ad-viewing behaviour should be considered in the context of the programmes that delivered the advertisements is more problematic. In other words the finding that two or more programmes with the same audience size could deliver radically different proportions of ad-viewing behaviours classed as being either eyes-on-screen, eyes-on-screen + behaviour, no eyes-on-screen, or not present suggest that programme ratings may be poor surrogate measures of advertising effect.

Lealand (1998) suggested that ratings have erroneously been used to make qualitative judgments about the "success or failure of single programmes, series (most specifically locally-made series), television channels and careers ... " (p.47). He argued that "... there is much to suggest that advertisers are being short-changed, particularly in respect of particulars about the television audience" (p. 41). He contended that "... watching television is accompanied by activities such as constant talk, eating, reading, household chores, quarrelling, clipping toe-nails, snoozing ... and so on." (p. 45). Ratings data therefore only presents part of the picture about ad viewing.

This presents the media planner with the need to adjust ratings data, in a way that accounts for these differences in viewing behaviour. The present study examined three methods by which this might be accomplished. However, the present study found viewing behaviour is not well predicted by programme categorisation. In some cases, the programme categories analysed in this research were found not to be different in terms of viewing behaviour. In other cases, the behaviour of the ad viewers was entirely contrary to those suggested by the programme categories. For example, while the categories of "demandingness" (developed by Barwise and Ehrenberg, 1988) were useful in showing the types of programmes viewers liked and watched repeatedly, this type of categorisation did not predict ad-viewing behaviour.

A major conclusion drawn from the analysis of programmes by genre is that presence is not a suitable and accurate surrogate measure for attention as there were wide variations in ad-viewing behaviour across the programmes. Relying solely on Peoplemeter or ratings data in media buying or as a measure of commitment to programmes and advertisements is not adequate.

Ad-viewing behaviour also tended to be different from how viewers rated the way programmes had an impact on them. Contrary to expectations, Low Impact programmes attracted as much attention as High Impact and High Cognitive programmes. These variations show the weaknesses of using subjective rating criteria to group programmes. In addition, there was also a low correspondence between programme genre and impact category suggesting considerable difficulties in developing an implementable solution.

For all three methods of classifying programmes, there were similarities and differences about how viewers behaved in different programmes. But three conclusions are common for all results. First, ad-viewing behaviour differs by programme type and thus needs to be considered separately for each particular programme type. In addition, the particular ad-viewing behaviour differs in an

unpredictable way for different programmes. Thus different conclusions might be reached if, for example, an advertiser sought to optimise in terms of *EOS only* than if the choice was made to optimise another characteristic of ad viewing, say *EOS + behavior*.

The last conclusion, and perhaps the underlying factor in the failure of the programme classifications, is that there may be many additional influences on ad-viewing behaviours that are not explicitly examined. These might reflect intrinsic programme or episode features, time of day, week, or in fact a number of highly specific situational influences.

This thesis has provided important information about how viewers behaved in their natural home environment, watching advertisements as they were screened, and highlighted some major considerations about the relationship between programmes and the actual attention paid to advertisements. Many of the conclusions outlined should be of concern to media planners, the advertising business in general, programme schedulers, programme funders and makers, and even those involved in providing programme ratings data.

This research has raised questions about traditional measures such as programme ratings and demonstrated that the size of an audience may not reflect how attentive that audience is. Even programmes of small audience sizes may have a high proportion of attentive ad viewers and therefore advertisements placed within such programmes may be equally (if not more) effective than programmes with a larger audience size.

As the television industry becomes more and more fragmented, advertisers who chose to segment their markets may need more and more specific information about the effectiveness of particular ad placements. Traditional TV ratings in use today cannot convey all the information needed by advertisers to know how audiences react to programmes and to their advertisements. In this respect, this

thesis has increased knowledge about general ad-viewing behaviour and the validity of using programme ratings in media scheduling decisions.

5.5 Limitations and Implications for Future Research

5.5.1 Limitations of the research

In any research, there are limitations. The major limitations of this research were associated with data collection and the possibility that moderating factors may have affected the viewing behaviour of the participants. Although this research collected data on the spontaneous responses of participants watching television in their own homes, it was dependent on how participants interacted with the recording devices. For example, while data collection depended on participants who agreed to watch television from the C-box, there was no way to monitor whether they always did. The research relied totally on the participants' goodwill.

Another factor is the question of audio attention versus visual attention during the viewing process. Viewers who rely on audio signals may reduce the amount of eyes-on-screen time they would spend looking at the television. A viewer may "learn" from the ad by listening without the watching. Such listening may already be familiar to viewers via repeated listening elsewhere. In other words, audio monitoring could diminish eyes-on-screen attention, especially when the advertisement was already familiar to the viewer, which would reduce the need to look at the screen.

This research did not measure whether visual attention was affected or assisted by audio sounds. However, unlike visual attention, which could be measured by eyes-on-screen, there was no simple analogous measurement for sound. This limitation may be eliminated if it later becomes possible to analyse audio data.

As different viewers watch television differently, the amount of time people spent watching television and the type of programme watched varied from

family to family. This meant, there may be variations in the data collected across all families. But the large number of ad-viewing opportunities recorded from each family, however, compensated for this limitation. In this research, analysis of these ad-viewing opportunities included how ad viewers behaved during "complete" ad pods over a consistent 8-day viewing period. The ideal situation would be to have many viewers (e.g. 50 participants) watching the same programmes at the same time. This would be possible if there were fifty C-Boxes operating at the same time in each of the 50 participants' homes. That would be highly expensive and perhaps probably impossible. Even if the cost could be justified and more data collected, there is no guarantee the results obtained would be significantly different from those collected in the present research.

In summary, the main limitations of this research resulted from difficulties associated with natural observational research. However, steps were taken where possible to eliminate or minimise the impact of any such limitation.

5.5.2 Implications for future research

Following the initial analysis of ad-viewing behaviour from this research, future work by this researcher will focus on collating information on the specific behaviour displayed during ad breaks—the proportion of time viewers were eating, playing games, talking, doing household chores, minding children, etc. This type of detail would supplement the existing information on the level of attention paid to ads, and more work could be done in studying the correlation between attention and other non-eyes-on-screen activities and how they differ in different programmes.

It is also important for advertisers to know the strategies viewers use to avoid looking at their ads. A detailed summary of these sorts of behaviour would

show any possible correlation with audience factors such as age, gender and other demographic characteristics. For example, gender has been found to be a factor in channel switching (Heeter & Greenberg, 1985; Danaher, 1995). Kneale (1988) also suggested there was a relationship between channel switching and income levels. More work in this area will provide up-to-date evidence to support these findings.

To know the direct effect of programming on commercial viewing, this study included only viewers who were watching the programme at least 30 seconds before the start of the ad breaks. It is highly likely that viewer's behaviour may change during the ad break from what they were doing during the programmes. But due to time and financial constraints, the behaviours during the programme segments prior to the start of the ad breaks or after were not analysed in this study. Earlier research such as Krugman et al. (1995) had found that the average mean attention to the screen was lower during commercials than during programmes. It may be worthwhile to investigate this further.

Furthermore, the viewing behaviour in different segments of a programme could be compared with that during the advertisement break immediately following the programme. For example, the behaviour of viewers watching the beginning section of the programme before the first ad break could be studied alongside the first ad break following it; programme viewing during the second segment of the programme (just preceding the ad break) could be studied in conjunction with ad behaviour for the second ad break; viewing behaviour during the third segment of the programme with the third ad break, and so on.

Other areas of future investigation could include studying viewing behaviour during different segments of the ad break. Questions to investigate might include, "Do viewers display different types of behaviour for first-position ads, ads in the middle or ads at the end?" and "Do viewing behaviour differ by the

position of the ad pod within a programme?" Such information would provide specific and detailed information about the actual "performance" of the ad.

Future research could also include studying the viewing behaviours of specific target audience groups such as children and teenagers. The effects of television on children have been of interest to psychologists for a long time, and academics have been worried about the adverse effects advertisements have on children. Work carried out by Anderson et al. (1985) has shown that parents underestimate the amount of time children spend watching television. They found parents reported that their children (age 2–5 years old) spent about 13.4 hours per week watching television, while the national ratings figure averaged 27.8 hours per week. In addition, little research has been carried out to investigate children's viewing behaviour during ad breaks.

Teenagers are an increasingly lucrative market for many advertisers. The New Zealand teenager has been dazzled by American popular dramas, such as *Beverly Hills 90210*, *Dawson Creek* and *Popular* that are specifically targeted at them. Other popular New Zealand dramas and youth programming include *Shortland Street*, *All about Eve*, *Havoc* and *Juice*. International youth-specific channels such as MTV, the international counterpart of New Zealand's Channel 4, have "relentlessly pursued the same audience" (Lealand & Martin, 2001). A study by Pollay (1996) suggested that teenagers were also more sensitive to advertising than adults and were susceptible to the image-based advertising often used in alcohol and tobacco products (Strasburger, 1995). On the other hand, the general public is mostly concerned with how commercials appeal to audiences "vulnerable" to their advertising messages, particularly messages in alcohol ads. Further research into teenage and child's responses to advertising could provide information on how best to use advertising to satisfy both the consumer and the public conscience.

In-depth investigative work on ad-viewing behaviour during individual programmes (not categorised in any particular programme groupings) could further the knowledge already obtained in this study. This could, however, be more time-consuming, and many different individual programmes might have to be studied separately. The methodology could include a random sample of daily programmes rated by individual viewers whose ad-viewing behaviour during these programmes could then be contrasted. In this way, there would be a more realistic correlation between the impact of the programme on the viewers and their actual ad-viewing behaviour during specific programmes.

Good advertising is a combination of the functions of the attributes of the ad, the attributes of the environment in which the ad appears, and the many variables that encompass the viewing experience. Although the set of variables surrounding the viewing environment is varied, it is not unlimited. Future work initiated by this research could look at extending the sample size, testing the correlation of eyes-on-screen and other dependent variables of viewing, studying the ad-viewing behaviour of specific target groups during specific ads and within specific programme segments.

6. REFERENCES

- Adspend 2001. Sector trends. (2002, March). *Admedia*, 60-61.
- Allen, C. (1965). Photographing the TV audience. *Journal of Advertising Research*, 5, 2-8.
- Anderson, D. (1985). Estimates of young children's time with television: A methodological comparison of parent reports with time-lapse video home observation. *Child Development*, 56, 1345-1357.
- Anderson, D., Lorch, E., Field, D., Collins, P. and Nathan, J. (1986). Television viewing at home: Age trends in visual attention and time with television. *Child Development*, 57, 1024-1033.
- Ang, I. (1991). *Desperately seeking the audience*. London: Routledge.
- Baron, R. (1995, March / April). The passive people meter: An agency perspective. *Journal of Advertising Research*, 2-4.
- Barwise, T. (1986). Repeat viewing of prime time TV series. *Journal of Advertising Research*, 26 (4), 9-14.
- Barwise, T. and Ehrenberg, A. (1987). The liking and viewing of regular TV series. *Journal of Consumer Research*, 14, 63-70.
- Barwise, T. and Ehrenberg, A. (1988). *Television and its audience*. London: Sage Publications.

- Barwise, T., Ehrenberg, A. and Goodhardt, G. (1979). Audience appreciation and audience size. *Journal of the Market Research Society*, 21 (4), 269-289.
- Bechtel, R., Achelpohl, C., and Akers, C. (1972). Correlates between observed behaviour and questionnaire responses on television viewing. In E.A. Rubinstein, G.A. Comstock, & J.P. Murray (Eds.), **Television and social behaviour, Vol.4**. USA: US Government Printing Office.
- Bennett, B. (1983, May). Rating levels and measuring techniques. *Admap*, 271-274.
- Billet, J. (1985, April). TV audience research: Future directions. *Admap*, 230-232.
- Bogart, L. (1989, February / March). What forces shape the future of advertising research? *Journal of Advertising Research*, 89.
- Bryant, J. and Comisky, P. (1978). The effects of positioning message within differently cognitively involving portions of a television segment on recall of the message. *Human Communication Research*, 5 (1), 63-75.
- Buck, S. (1983, December). TV audience research - present and future. *Admap*, 636-640.
- Clancey, M. (1994). The television audience examined. *Journal of Advertising Research*, 34 (4), 2-11.
- Clancy, K. and Lloyd, D. (1999). *Uncover the hidden power of television programming ... and get the most from your advertising budget*. California: Sage Publications, Inc.

- Coen, B. (2001, December). *Bob Coen's Insider Report*. Retrieved February 21, 2002, from <http://www.mccann.com/insight/bobcoen.html>
- Collett, P. and Lamb, R. (1986). *Watching people watch television*. Report for the Independent Broadcasting Authority. Oxford, UK: University of Oxford Press.
- Colmar Brunton Research (1992). *Viewer behaviour and recall during TV breaks*. Unpublished report to ANZA, Auckland.
- Cook, B. (1995, March / April). Commercial television: Dead or alive? A status report on Nielsen's passive people-meter. *Journal of Advertising Research*, 5-10.
- Danaher, P. (1995). What happens to television ratings during commercial breaks? *Journal of Advertising Research*, 37-47.
- Danaher, P. and Beed, T. (1992, March). Peplemeters - Do they really measure what people watch? *Admap*, 42-44.
- Danaher, P. and Beed, T. (1993). A coincidental survey of people meter panellists. *Journal of Advertising Research*, 33, 86-92.
- Danaher, P. and Lawrie, J. (1998). Behavioural measures of television audience appreciation. *Journal of Advertising Research*, 54-65.
- den Boon, A. (1996). From people meter to decoder. Revolution on the TV screen. *European Society for Opinion and Marketing Research*. Retrieved February, 2002, from <http://www.warc.com/print/9517p.asp>

- Ehrenberg, A. and Goodhart, G. (1981). Attitudes to episodes and programmes. *Journal of Market Research Society*, 23 (4), 189-208.
- Ehrenberg, A. and Twyman, W. (1967, November). On measuring television audiences. *Journal of the Royal Statistical Society, Series A*, 130, 1-59.
- Ehron, E. and Gray, S. (2001, February). Why we can't afford to measure viewers. *Journal of Advertising Research*. Retrieved February 27, 2002, from <http://www.warc.com/print/49904p.asp>
- Gane, R. (1993, February). A world of peplemeters. *Admap*, 25-30.
- Gane, R. (1994). Television audience measurement systems in Europe. In Raymond Kent (Ed.), *Measuring media audiences* (p. 22-42). Routledge: London.
- Gill, J. (2000, December). Managing the capture of individuals viewing within a peplemeter service. *Journal of the Market Research Society*. Retrieved March 8, 2002, from <http://www.WARC.com/print/348435p.asp>
- Goldberg, M. and Gorn, G. (1987). Happy and sad TV programs: How they affect reactions to commercials. *Journal of Consumer Research*, 14, 387-403.
- Green, A. (2001, October). The East is reddening: TV in China. *Admap*, 48-50.
- Gunter, B., Furnham, A. and Beeson, C. (1997). Recall of television advertisements as a function of programme evaluation. *The Journal of Psychology*, 13 (15), 541-553.

- Gunter, B., Furnham, A. and Lineton, Z. (1995). Watching people watching television : What goes on in front of the TV set? *Journal of Educational Television*, *21* (3), 165-191.
- Heeter, C. and Greenberg, B. (1985). Profiling the Zappers. *Journal of Advertising Research*, *25* (2), 9-12.
- Hoffman, D. (1991). *Multivariate analysis of audience attitudes and behaviours during television viewing (ratings, advertising media)*. Unpublished doctoral thesis. University of North Carolina, Chapel Hill.
- Hoffman, D. and Batra, R. (1991). Viewer response to programs: Dimensionality and concurrent behaviour. *Journal of Advertising Research*, *31* (4), 46-56.
- Kamins, M., Marks, L. and Skinner, D. (1991). Television commercial evaluation in the context of programme induced mood: Congruency versus consistency effects. *Journal of Advertising*, *20* (2), 1-14.
- Kennedy, J. (1971). How programme environment affects commercials. *Journal of Advertising Research*, *11* (1), 33-38.
- Kent, R. (Ed.). (1994). *Measuring media audiences*. London: Routledge.
- Krugman, D., Cameron, G. and White, C. (1995). Visual attention to programming and commercials: The use of in-home observations. *Journal of Advertising Research*, *24* (1), 1-12
- Krugman, D. and Shamp, S. (1992). Observations of audience behaviour during television commercials. *Proceedings of the 1992 conference of the American Academy of Advertising*, 156-160.

- Krugman, H. (1983). Television program interest and commercial interruption. *Journal of Advertising Research*, 23 (1), 21-23.
- Lang, A. (Ed.). (1994). *Measuring psychological responses to media messages*. New Jersey: Lawrence Erlbaum Associates.
- Lealand, G. (1998). Ratings and more damn ratings: Measuring television viewing in New Zealand. *New Zealand Journal of Media Studies*, 5 (1), 36-50.
- Lealand, G. and Martin, H. (2001). *It's all done with mirrors: About television*. Palmerston North: Dunmore Press.
- Lee, B. and Lee, R. (1995). How and why people watch TV: Implication for the future of interactive television. *Journal of Advertising Research*, 9-18.
- Lind, H. (1984, January). What counts in the media? *Marketing*, 45-47.
- Lindsay, S. (2001, July). How well do TV ratings really rate? *Onfilm*, 15.
- Lloyd, D. and Clancy, K. (1991). Television programme involvement and advertising response: Some unsettling implication for copy research. *Journal of Consumer Marketing*, 8 (4), 61-74.
- Lloyd, D. and Clancy, K. (1991). CPMs versus CPIMs: Implications for media planning. *Journal of Advertising Research*, 34-44.
- Lord, K. and Burnkart, R. (1993). Attention versus distraction: The Interactive effect of programme involvement and attentional devices on commercial processing. *Journal of Advertising*, 22 (1), 47-60.

- Lull, J. (1991). *Inside family viewing. Ethnographic research on television's audiences*. London: Routledge.
- Lynch, K. and Stipp, H. (1999). Examination of qualitative viewing factors for optimal advertising strategies. *Journal of Advertising Research*, 7-16.
- Market Research News. (2002, January). *Final Test for Arbitron meter*. Retrieved February, 15, 2002, from <http://www.mrnews.com/12ts3ny/news/ppm-test.html>
- McDonald, C. (2000, May). *Television Audience Measurement*. Retrieved February 27, 2002, from <http://www.warc.com/fulltext/BestPractice/13902.htm>
- Merz, T., Schumacher, G. and Sudholt, T. (1993, November). How accurate are peplemeters? *Admap*, 43-46.
- Milavsky, J. (1992). How good is the peplemeter system? *Public Opinion Quarterly*, 56 (1), 102-115.
- Monten, R. (1989). You count heads - But can you describe with meters ? ESOMAR. *Seminar on broadcasting research: Experiences and strategies*, 179-185.
- Moriarty, S. and Everett, S. (1994). Commercial breaks: A viewing behaviour study. *Journalism Quarterly*, 71 (2), 346-355.
- Morley, D. (1990). Behind the ratings: The politics of audience research. In J. Willis & T. Wollen (Eds.), *The neglected audience*. London: BFI Publishing.

- Murphy, J., Cunningham, I. and Wilcox, G. (1979). The impact of program environment on recall of humorous television commercials. *Journal of Advertising Research*, 8 (2), 17-21.
- Murray, J., Lastovicka, J. and Singh, S. (1992). Feeling and liking responses to television programmes: An examination of two explanations for media-context effects. *Journal of Consumer Research*, 18, 441-451.
- Norris, C. and Colman, A. (1994, January). Putting ads in context. How television programmes affect viewers' reactions to ads. *Admap*. Retrieved March 7, 2002, from <http://www.WARC.com/print/3363p.asp>
- NZOnAir, (2000). *Local New Zealand Content 2000*. New Zealand: NZOnAir.
- Perry, S., Jenzowsky, S., King, C., Yi, H., Hester, J. and Gartenschlaeger, J. (1997). Using humorous programs as a vehicle for humorous commercials. *Journal of Communication*, 47 (1).
- Pollay, R., Siddarth, S., Siegel, M. and Haddix, A. (1996). The last straw? Cigarette advertising and realized market shares among youths and adults. 1979-1993. *Journal of Marketing*, 60 (2), 1-16.
- Poltrack, D. (1998). Media planners look beyond the numbers. *Admap*. Retrieved March 4, 2002, from <http://www.warc.com/print/6139p.asp>
- Purdye, K. and Harvey, B. (1994, November). TV audience measurement around the world. *Marketing Research Today*, 223-230.
- Rawlings, T. (1983, May). The hi-tech alternatives. *Admap*, 268-271.

- Reeves, B., Thorson, E. and Schleuder, J. (1986). Attention to television: Psychological theories and chronometric measures. In J. Bryant & D. Zillman (Eds.), *Perspectives on media effects* (p. 251 – 279). New Jersey: Lawrence Erlbaum Associates.
- Rubens, W. (1978, February). A guide to TV ratings. *Journal of Advertising Research, 18 (1)*, 11-18.
- Schwerin, H. and Newell, H. (1980). *Persuasion in Marketing*. Wiley: New York.
- Schumann, D. and Thorson, E. (1990). The influence of viewing context on commercial effectiveness: A selection-processing model. *Current Issues and Research in Advertising, 12*, 1-24.
- Sharot, T. (1994). Measuring television audiences in the UK. In Raymond Kent, (Ed.), *Measuring media audiences* (p. 42-87). Routledge: London.
- Soldow, G. and Principe, V. (1981). Response to commercials as a function of program context. *Journal of Advertising Research, 21 (2)*, 59-65.
- Steiner, G. (1966). The people look at commercials: A study of audience behaviour. *Journal of Business, 39 (2)*, 272-304.
- Strasburg, V. (1995). *Adolescents and the media: Medical and psychological impact*. Sage Publications: Thurmond Oaks.
- Svennevig, M. and Wynberg, R. (1986, May). Viewing is viewing is viewing ... or is it ? A broader approach to television research. *Admap, 267-275*.

- Svennevig, M. (1987, January). Viewers are people too: A fresh look at the audience. *Research Philosophy*, ESOMAR, 75-96.
- Syfret, T. (1995, November). Measuring television audiences beyond 2001. *Admap*, 41-44.
- Thorson, E. (1994). Using eyes on screen as a measure of attention to television. In A. Lang (Ed.), *Measuring psychological responses to media messages*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Thorson, E. and Reeves, B. (1986). Effects of overtime measures of viewer liking and activity during programs and commercials on memory for commercials. In R. Lutz (Ed.), *Advances in consumer research (Vol. 13)*. Provo, UT: Association for Consumer Research.
- Zwaga, W. (1992a). *Families in front of the screen*. Unpublished doctoral dissertation, Massey University, New Zealand.
- Zwaga, W. (1992b). Delivering television audiences to the advertisers? Impressions from the living room. *Marketing Bulletin*, 3, 9-17.

Appendices

Appendix A: The C-Box



Appendix B: Participant's Programme Evaluation Questionnaire

Date: _____

Name of programme: _____

1. Did you look forward to the programme ?

Yes No

2. How did this programme compare with the best of its kind ?

1 2 3
worse better than most

3. How much did you enjoy this programme ?

1 2 3
not much very much

4. Did this programme match up to your expectations ?

1 2 3
better than expected not so good

5. Did the programme touch your feelings ?

Yes No

6. Did you watch this programme as a way to pass time ?

Yes No

7. How would you grade this programme ?

1 2 3 4 5 6 7 8 9
Bad Good

8. Did you learn anything from the programme ?

Yes No

9. Was the programme unusual or different ?

Yes No

10. Did you plan ahead to watch this programme ?

Yes No

11. Would you be upset if, given that the programme is on tomorrow, it was missed ?

Yes No

12. I usually give my full attention to the programme so that I could get more out of it.

1 2 3 4 5
agree disagree

13. I really get involved in the programme and wished it had lasted longer.

1 2 3 4 5
agree disagree

14. At times, the programme made me feel the same things as the characters in it.

1 2 3 4 5
agree disagree

15. There were parts of the show that really touched my feelings.

1 2 3 4 5
agree disagree

16. The programme contained valuable and vital information I am interested in and means a lot to me.

1 2 3 4 5
agree disagree

17. The funny parts of the programme were very humorous and enjoyable.

1 2 3 4 5
agree disagree

18. I watched this programme mainly to be entertained.

1 2 3 4 5
agree disagree

Appendix C: Participant's Information Sheet

MASSEY UNIVERSITY
DEPARTMENT OF MARKETING
PALMERSTON NORTH

TELEVISION VIEWING RESEARCH INFORMATION SHEET

The research project "Television Viewing Research" is designed to study the ways in which families use television in the context of their everyday life, with a particular emphasis on viewing behaviour during TV commercials. While we know, for instance, that most New Zealand families own one or several television sets, we have presently little information about the role of television within the family. The present study is designed to provide information on this topic, paying particular attention to viewing behaviour and to what people do during television programmes and commercials.

The researchers involved in this project are Dr Mike Brennan, Dr Graeme Bassett and Ms Maureen Syn. Mike Brennan is a Senior Lecturer in the Department of Marketing at Massey University. Graeme Bassett is Director of Media Studies at Massey University. Maureen Syn is a Lecturer at International Pacific College, currently undertaking this research for her PhD. If you have any questions about the project, or the consent forms, please feel free to contact Dr Mike Brennan 359 0705 (Home) or 350 5576 (Work) or Ms Maureen Syn 357 1786 (Home) or 354 0922 (Work) for further information.

Should you agree to take part in the research the following procedures will be involved. Firstly, a cabinet consisting of a television set, two video recorders and a video camera will be placed in your home for fourteen days. This cabinet will be wired up so that when the television set is switched on, the video recorders, through the camera, will start recording the activities of whoever is present in the living room as well as the programme being watched. An

additional video recorder will be included in the cabinet for those families who normally use a video recorder.

Secondly, in the event of your agreeing to take part, one or two interview sessions will take place with you and your family. These interviews will be sound recorded. Both stages in the research will be arranged at times mutually agreed to by the participants and the researchers, with the understanding that the interviews ideally should take place no later than one week after the cabinet has been removed from your home. Thirdly, a nominated member of your household may be asked to keep a diary of their TV viewing activity.

Participation in the "Television Viewing Research" project is entirely voluntary. Participants are free to withdraw from the research at any time without threat of reprisal or recrimination from the researchers. Practically, this means that when one family member wants to withdraw from the research, the family as a whole will discontinue participation. Participants have the right to have any recorded material erased, if they feel that nobody should have access to what has been recorded on tape. If this should occur, participants consent would be re-canvassed for further involvement. The names of the participants will not be released and no identifiable material shall be shown publicly nor published in whatever possible form, except with the explicit written permission of the participants involved.

If you agree to take part, please read this Information Sheet again carefully before signing the consent form. (Please note that the consent form should be signed by all family members aged ten and over.) It may be a good idea to read this sheet together and discuss it with all members of your family.

It would also be important to inform visitors of the nature of the cabinet and say that you are taking part in a Massey University research project. If visitors arrive during the recording of the video, the machine should be turned off, the

research project explained to them and their agreement to being included in the project sought, before the machine is turned on again. Training will be provided in the use of the equipment, to enable you to stop the recording, or erase tapes, if required.

All information will be treated in strictest confidence. The video and audio recordings will be identified by a code number only, not by family name. They will be securely stored during the course of the research and will only be viewed for research purposes by Dr Mike Brennan, Dr Graeme Bassett or Ms Maureen Syn and their research assistants who will sign an agreement that they will maintain the strictest level of confidentiality. At the completion of the study, the recordings will be securely stored in an archive that can only be accessed by Dr Brennan, Dr Bassett or Ms Syn. The researchers may wish to use the recordings for further analysis at some later date, involving other researchers in their team. Should this situation arise, they would first explain to you, both verbally and in writing, exactly what they want to do, and will not proceed without your written consent.

The cabinet with the television set, the video recorders and the video camera are the property of Massey University. The cabinet will be insured by Massey University for the period the cabinet is in the participants' residence. The participants shall not be liable for any breakdown or accidental damage that may eventuate.

We would greatly appreciate your assistance with this project. Please feel free to ask at any time if you have any further questions.

Dr Mike Brennan	Dr Graeme Bassett	Maureen Syn
Principal Researcher	Associate Researcher	Researcher
359 0705 (home)	357 3584 (home)	357 1786 (home)
350 5576 (work)	350 5356 (work)	354 0922 (work)

Appendix D: Participant's Consent Form

**TELEVISION VIEWING RESEARCH
CONSENT FORM**

I have read and understood the TELEVISION VIEWING RESEARCH INFORMATION SHEET and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand I may ask further questions at any time.

I agree to allow the video and audio recording of my television viewing on the understanding that the recordings will be stored securely and treated in strictest confidence and will not be used for purposes other than this study without my permission.

I understand I have the right to withdraw from the study at any time, and have the right to stop the recording at any time and to request that recorded material be erased.

I agree to participate in this study under the conditions set out in the INFORMATION SHEET and give my permission for the recordings and information gathered to be used for the purposes of this research then archived.

Name: _____ Signature _____

Name: _____ Signature _____

Name: _____ Signature _____

Name: _____ Signature _____

Date: _____

Appendix E: Summary of participant families

Family	Family members & position	Age	Occupation	Ethnicity	Household income
1	Father Mother Son Son Daughter Son	36 33 11 8 5 1	Marketing Manager Home executive Child Child Child Child	European European European European European European	\$50,000 to \$59,999
2	Father Mother Daughter Son Son	41 38 7 5 3	Land valuer Home executive Child Child Child	European European European European European	\$90,000+
3	Father Mother Son Son Daughter	41 40 12 9 7	Teacher Teacher Child Child Child	European European European European European	\$50,000 to \$59,999
4	Father Mother Son Boarder	51 50 23 23	Surveyor Bank teller Student Student	European European European Japanese	Unknown
5	Flatmate (female) Flatmate (female)	32 26	Bank officer Bank teller	European European	\$80,000 to \$89,999
6	Husband Wife	50 49	Developer Receptionist (part-time)	European European	\$90,000+
7	Father Mother Son Son Daughter	45 39 16 13 11	Lecturer Part-time youth worker Student Child Child	European European European European European	Unknown
8	Husband Wife Boarder	Unknown 26 25	Sales executive Bank officer Bank teller	European European European	Unknown
9	Husband Wife	41 50	Lecturer Activity Manager	European European	\$90,000+
10	Single female	37	Part-time caregiver	European	Less than \$20,000
11	Father Mother Son Daughter	51 44 13 9	Community worker Administrator Child Child	European European European European	\$90,000+
12	Father Mother Son Daughter	36 35 5 2	Lecturer Administrative assistant Child Child	European European European European	\$70,000 to \$79,999
13	Father Mother Daughter Son	39 36 5 2	Technical Manager Cosmetics Distributor Child Child	European European European European	Unknown
14	Father Mother Daughter	Unknown Unknown Unknown	Lecturer Teacher Child	Indian Indian Indian	Unknown

Appendix F: Rationale for 30-second criterion for ad pod selection

This section explains the rationale for selecting that the viewer must be continuously tuned in to the programme 30 seconds before the start of the ad break.

Involvement with programmes

The viewing behaviour during the ad breaks was used to study the relationship between programme and viewer involvement with the ads during the programmes. To make sure that the programme was relevant to the commercial break being studied, it was important to be sure that the viewer had spent sufficient time "watching" the programme to be judged as "involved" with it. It was possible that viewers could have returned from being away from the programme or had just tuned in to the ad break from surfing between channels. This situation would diminish the value of the analysis of the ad-viewing behaviours. In other words, there must be proof that the viewer was in the room and tuned to the programme so that comparison between programme involvement and ad involvement could be meaningful.

However, there are several different criteria pertaining to the length of programme watched that could be used to select the ad breaks. Three criteria were chosen. The first was that the viewer was tuned in to the programme for at least 4 minutes before the start of the ad break. The second was that the viewer was there for at least 80% of the 4 minutes preceding the ad break. The third criterion was that the viewer was there continuously for 30 seconds or more before the start of the ad break.

Programme type comparison

Different researchers have used different methods to classify programmes in terms of "involvement". The methods used in three different studies on evaluating programme types have been adapted in this study. Firstly, Barwise and Ehrenberg (1988) developed two major types of programme based on the level of content the programme contained. They termed them as "*demanding*" or "*less-demanding*" (Entertainment programmes).

Secondly, Danaher (1995) and Danaher and Lawrie (1998) measured viewers' "commitment" by the length of time viewers were present in the room during the different programmes. According to them, "commitment" refers to the proportion of time their viewers were present for the ads out of the potential time the ads were screened. Their classifications include broad generalisations based on common genre themes.

Lastly, Hoffman and Batra (1991) used three measures of "programme impact" that according to the researchers measures cognitive and affective reactions to the programmes. They grouped programmes into: high cognitive impact programmes; high affective programmes and low impact programmes.

Behavioural measures by each criterion

In the present study, the respondents' viewing behaviours during the ad break were examined for the above three different ways of classifying programme types to determine whether there are reliable ways to group programmes according to their involvement or impact value on the viewers. For each programme type, the data was aggregated according to the selected criteria and means percent scores for four behavioural measures were obtained. The measures were eyes-on-screen only (*EOS only*), meaning the person was seen looking at the screen; eyes-on-screen and engaging in one or more simultaneous

activities (*EOS + behaviour*); not looking at the screen (*No EOS*); and left the room (*not present*). To test whether the three criteria used produced different results, the overall total means of the four categories of viewing behaviours were compared.

As the number of observations for the programme types using Barwise and Ehrenberg (1988) classification (i.e. all programme recorded) is exactly the same as that using Danaher (1995) classification, both classification methods have similar overall means. For this reason, only a single analysis was performed (see Table A1)

Table A1: Comparison of ad-viewing behaviour

Criteria	No. of ads observed						Total %
			<i>Present</i>		<i>Not present</i>		
			<i>EOS only</i>	<i>EOS + behaviour</i>	<i>No EOS</i>		
30 sec preceding	6017	Mean	14	11	52	23	100
4 min preceding	4661	Mean	15	12	53	20	100
80% of 4 min present	5212	Mean	14	12	53	21	100
Average	5297		14	12	53	21	

Table A2 presents the findings of the viewing behaviours using Hoffman and Batra (1991) classification. The number of ad observations (i.e. all programmes rated) differed from Barwise and Ehrenberg (1988) and Danaher (1995) because the two methods used different measures to rate programme

involvement. Some programmes did not receive any ratings as there were instances where viewers could not recall the programmes watched and could not therefore comment on how those programmes had impacted them. In other words, Table A2 describes a slightly smaller sub-set of those ads described in Table A1.

Table A2: Comparison of ad-viewing behaviour using data from viewers' personal evaluation of programmes (Hoffman & Batra, 1991)

Criteria	No. of ads observed					<i>Not</i>	<i>Total</i> %
			<i>Present</i>		<i>present</i>		
			<i>EOS</i> <i>only</i> %	<i>EOS +</i> <i>behaviour</i> %	<i>No</i> <i>EOS</i> %		
30 sec preceding	5373	Mean	12	12	53	22	100
4 min preceding	4260	Mean	13	12	53	22	100
80% of 4 min present	4763	Mean	13	13	53	22	100
Average	4799		13	12	53	22	

From the above tables, it could easily be seen that the differences in the mean scores for the three selection criteria used were quite small. In all cases the difference is no larger than 1%. The difference between the three sets of results is not far from each other for the Ehrenberg and Danaher classifications ($\chi^2 = 0.1229$, $df = 6$, $p > 0.5$) or the Hoffman & Batra classifications ($\chi^2 = 0.0529$, $df = 6$, $p > 0.5$). Rather than duplicating similar results, the criterion, the viewer

was present in the room 30 seconds before the start of the ad break, was used to select the observations that were analysed.

Appendix G: Examples of ad-viewing behaviour

EOS only

0 Looking at the screen only

EOS + behaviour

111 Eating / drinking with EOS

113 Eating / drinking and other activity (with EOS)

121 Feeding (baby or pet) with EOS

123 Feeding and other activity (with EOS)

131 Knitting / sewing with EOS

133 Knitting / sewing and other activity (with EOS)

141 Smoking with EOS

143 Smoking and other activity (with EOS)

211 Attending to baby or pet with EOS

213 Attending to baby or pet and other activity (with EOS)

221 Cleaning / tidying / walking around packing with EOS

223 Cleaning / tidying / walking around packing and other activity (with EOS)

241 Games / playing with EOS

243 Games / playing and other activity (with EOS)

251 Muting the sound but with EOS

281 Talking with EOS

283 Talking and other activity (with EOS)

No EOS

- 112 Eating / drinking with no EOS
- 114 Eating / drinking and other activity (with no EOS)
- 122 Feeding (baby or pet) with no EOS
- 124 Feeding and other activity (with no EOS)
- 132 Knitting / sewing with no EOS
- 134 Knitting / sewing and other activity (with no EOS)
- 144 Smoking with no EOS
- 146 Smoking and other activity (with no EOS)

- 212 Attending to baby or pet with no EOS
- 214 Attending to baby or pet and other activity (with no EOS)
- 222 Cleaning / tidying / walking around packing with no EOS
- 224 Cleaning / tidying / walking around packing and other activity (with no EOS)
- 242 Games / playing with no EOS
- 244 Games / playing and other activity (with no EOS)
- 252 Muting the sound but with no EOS
- 260 Reading
- 270 Sleeping
- 282 Talking with no EOS
- 284 Talking and other activity (with no EOS)
- 290 Switching channel

Not present

- 310 Left / leaving the room
- 320 Not in room / not in view

Glossary

ADRATIO	The number of people who stayed for the ad break out of the number of people who were present for the programme.
Ad avoidance	When viewers deliberately choose not to watch the advertisements screened.
Ad break	When programme broadcasts are interrupted and at least more than one paid advertisement is being screened.
Ad-viewing behaviour	The viewing behaviour of viewers in the room with the television set on during the ad break.
Ad pod	A set of advertisements (between two to fourteen) that interrupted the programme.
Ad position	This refers to the position of each individual advertisement found within an ad break.
Ad pod position	This refers to the position of the ad break as it was scheduled within a programme.
Audience ratings, Ratings	The percentage of a particular population watching a television programme, commercial or channel at a particular time. Usually refers to the number of people who are in the room with the television set switched on.
C-Box	A cabinet that housed a hidden camera and video-recorders programmed and connected to a TV that sits on top of the cabinet.
Cable television	A system by which television signals are transmitted via over or underground cables, rather than over the air.
Chat show, talk show	A show where a host and guests are chatting in front of a studio audience.

	The programme is sometimes interspersed with music and a variety show.
Children's programmes	Programmes such as cartoons which are targeted at younger audiences.
Coincidental surveys	Survey interviews (face-to-face or phone) asking viewers whether and what channels the informant and / or others in the household are watching at the moment the interviewer calls.
Comedy	Light hearted programmes designed for light entertainment, usually based on the humour of a character or situation.
Competitive entertainment	These programmes have broad competitive formats and serve to entertain rather than provide information.
Current affairs programme	Programmes that deal with topical issues of current and public concerns (political, economic, social etc.)
Demandingness	The degree in which a programme requires intellectual and / or emotional effort and demands higher involvement from its viewers.
Diary methods	People recording their viewing habits manually into a 7-day diary.
Drama	These are programmes that follow drama production techniques and have a continuous story line. They include stories based on legal, crime and lifestyle vignettes. They may contain humorous storyline or funny idiosyncratic characters such as "Ally McBeal".
Eyes-on-screen (EOS)	This refers to the action where viewers' eyes are fixated at the screen.
<i>EOS only</i>	Observations where viewers were looking at the screen as a sole activity.

<i>EOS + behaviours</i>	Observations where viewers have eye-on-screen and also doing something else at the same time.
<i>EOS only: Presence</i>	The proportion of time when viewers were looking at the screen during the time they were present in the room.
Game show / Quiz	Programmes following a game format and usually involve celebrity or audience participation.
High Affective programmes	Programmes rated by viewers as those which had touched their feelings.
High Cognitive programmes	Programmes rated by viewers as those they had learnt something from.
High Impact programmes	Programmes rated by viewers as those which had both touched their feelings and which they had learnt something from.
Information programmes	News magazines, current affairs, documentaries and feature programmes. Also included general interest programmes such as religion, motoring, gardening, cooking, consumer programmes etc.
Informative interest programmes	These programmes follow a magazine format with an information flavour. The content may serve special interest topics such as fishing, gardening, religion, arts and culture programmes, leisure and lifestyle series.
Involvement	Broadly, the degree of viewers' commitment or attention to what they are viewing.
Light drama	These are situation comedies, light plays and drama (including made for TV mini-series and one-off programmes).

Light Entertainment	Variety and music programmes. Includes quiz shows, panel games, chat shows and other "personality" programmes.
Light viewers	People who tend to watch relatively little television (less than 10 or 15 hours weekly).
Low Impact programmes	Programmes rated by viewers that neither touched their feelings nor did they learnt anything.
Movies	These are one-off drama programmes that follow a movie production format.
Musicals	These programmes include one-off special variety events or regular weekly music charts.
News, documentaries	Programmes that include New Zealand and overseas news current affairs and current affairs programmes, one-off advertised events or are regularly screened.
Opportunities to see, OTS	The number of times when a viewer is estimated to be present with the television set tuned to the relevant channel while the commercials of a particular advertising campaign are being shown.
P80+	A measure that stipulated that a viewer was present for at least 80% of the proportion of time viewers were present during the ad break out of the total time the programme was aired.
Panel	In audience research, this refers to a group of households who take part continuously (or over a set period of time) in survey measurements of television audiences.

Peak viewing	The time of the day when the largest number of people tend to be viewing television; from 6.30p.m. to 10 p.m. (prime time) in most countries.
Peoplemeters	Electronic push-button ways of recording the presence of viewers in a household when the television set was switched on.
PMV	The number of minutes viewed (by being there) out of the number of minutes the programme was aired.
Presence	This refers to the viewer being in the room when the television set was switched on regardless if viewer was watching (looking at) the programme.
Propview	The proportion of ad minutes viewed during the ad break out of potential viewing time during the ad break.
Prime time	This refers to the hours when the television ratings are highest. Used with very specific time limits in connection with costs of airtime for commercials.
Repeat viewing	This happens when a viewer of one episode of a regular programme watches a following episode.
Sitcom, situation comedy	Comedies based on the humour of character and situation and where the plot usually gets resolved at the end of the episode.
Soaps, mini series and serials	Narrative drama programmes that follow a continuous story line over successive weeks and sequence episodes. They usually contain a continuing human interest and topical issues.
Sports programmes	These are mainly New Zealand produced sporting programmes but may contain a high level of overseas content.

Time-shift

Using a video recorder to tape a programme off-air for playback at a time convenient to the viewer at a later time.

Total EOS

The proportion of total time viewers had their eyes on the screen (the sum of *EOS only* and *EOS + behaviour*).

Total EOS:Presence

The proportion of total eyes-on-screen time out of the time viewers were present during the ad break.