International Journal of Business and Social Research
Volume 04, Issue 12, 2014

Measuring Emotional Responses to TV Commercials: The Warmth Monitor Modernized

Marc Roy[1]

ABSTRACT

Recently there has been a lot of interest in measuring emotional responses to advertising. This study focuses on the measurement of a specific emotional response to television advertising; warmth. Nearly thirty years ago, (Aaker, Stayman and Hagerty, 1986) developed a procedure they called the Warmth Monitor; "paper and pencil" self-report process recording method. The Warmth Monitor has been used in a large number of empirical studies in marketing since, but the most recent versions of the procedure are computerized. The two methods of administering the Warmth Monitor are compared in this research.

Keywords: Advertising effectiveness, commercial breaks, emotional response, felt mood, warmth monitor.
JEL Codes: M3, M37.
Available Online: 29th December, 2014.
MIR Centre for Socio-Economic Research, USA.

1.0 INTRODUCTION

In a study focussing on commercial breaks placement, using students enrolled in business administration at a Canadian University, (Roy, 2013) showed that the way a commercial interrupts a TV program will influence a viewer's felt mood that will, in turn, affect his reactions to an advertisement. For this experiment, a TV commercial with the following features was needed:

- A very low information content level;
- An advertised brand not available in Canada;
- A commercial never shown on Canadian TV;
- A very low product category implication from the participants;
- A TV commercial content conductive to experiencing a warmth feeling.

[1] Professeur, Sciences de Gestion, Université du Québec à Rimouski, Canada, E-mail: marc_Roy@uqar.ca.

To measure the feeling of warmth created by a pair of TV commercials satisfying the above criteria, two versions of the Warmth Monitor were employed: the "paper and pencil" procedure and a computerized one. These 2 advertisements were selected by carrying out a pre-test that was performed on a comparable student sample (n = 60) to the one used in the main study. The participant rated 25 commercials using a single post exposure self-report scale; a measurement procedure typical to many researches on emotional ad response (e.g. Plutchik, 1980; Richins, 1997; Lioret, Segura and Gonzzalez-Roma, 2003).

Consistent with previous research, the hypotheses of the study are:

H1: Average scores of the feeling of warmth created by TV commercials
The Warmth Monitor's reliability has been established in a number of past studies (e.g. Aaker, Stayman and Hagerty, 1986; VandenAbeele and Maclachlan, 1994; Baumgartner, Sujan and Padgett, 1997; Derbais and Poncin, 2005). Consequently, the average warmth scores obtained from both Warmth Monitor procedures will not be significantly different.

H2: Variance of warmth scores from Warmth Monitor procedures
Since the computerised version of the Warmth Monitor only requires the participants to control a cursor's horizontal position to indicate to what level they feel more or less warmth, their task will be simpler than for those using the "paper and pencil" recording method. Consequently, the measures obtained using the computerised version of the Warmth Monitor will vary less between the participants than for those using the "paper and pencil" recording method. Thus, the variance of the scores compiled from the computerised procedure will be significantly inferior to the variance of those coming from the "paper and pencil" method.

The focus of this study is not on examining the effects of the feeling of warmth created by TV commercials, but on comparing the two procedures used to measure them. The remainder of the paper reports on the laboratory study conducted to test the two versions of the Warmth Monitor. It is organized as follows: First, the conceptual background is presented. Next, the necessary analyses for hypotheses testing are conducted and finally a discussion of the findings is offered.

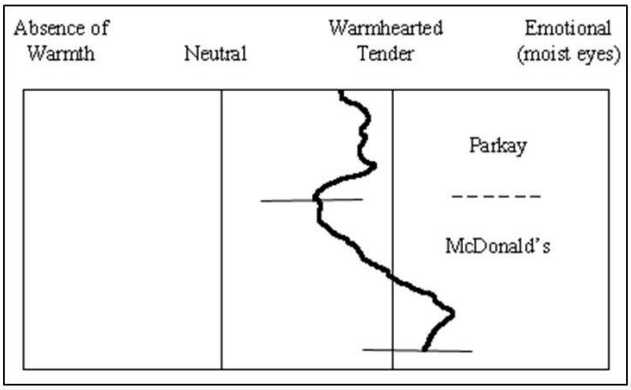
2.0 LITERATURE REVIEW

A number of researchers have studied the effects of program environment on felt mood and, subsequently, on television advertisement effectiveness (e.g. Batra and Ray, 1986; Edell and Burke, 1987; Coulter, K. S., 1998; Medianu, S., 2007; Owolabi, A. B., 2009). Traditionally, post exposure verbal responses are most often used to rate affective reactions elicited by commercials. Such scales are easy to administer but they require that the respondent provides a summing of a viewing experience that usually lasts 30 or 60 seconds. The changeability of affective reactions means that a single measure is purporting to cover the entire commercial may be very misleading (Fenwick and Rice, 1991, p.24).

In a very famous article, (Aaker, Stayman and Hagerty, 1986) have studied the impact of the feeling of warmth created by a commercial; using a measurement procedure they called the Warmth Monitor. It was a "paper and pencil" self-report process recording method employed to obtain a real-time measure of an emotional response (warmth) to TV ads.

The Warmth Monitor required that the participant drew a vertical line down a paper sheet while viewing a commercial, moving the pencil left or right to reflect how warm his feelings were at any given time. Many researchers agree that such a procedure is more efficient than post exposure verbal responses for measuring emotional responses to advertising. "Process tracing methods are a better solution to the extent that they offer the opportunity to examine patterns of response to transient external stimuli over time" (Vanden Abeele and Maclachlan, 1994, p.597).

Figure 1: The Warmth Monitor Aaker, Stayman and Hagerty



Comparing the Warmth Monitor to other methods, (Vanden Abeel and Mac Lachlan, 1994) found that the Warmth Monitor provided a reliable measure of warmth, but recommended that additional research be done to improve its validity. (Derbais and Poncin, 2005) stated that recent versions of these types of measuring scales are computerized. Affective reactions are now registered using a computer mouse. In this particular case the cursor goes down the screen at a constant speed and the participant controls its horizontal position by moving the mouse left and right. Thus, the participant indicates to what level he feels more or less warmth at every moment the commercial is being viewed. For (Derbais and Poncin, 2005), the computerized version of the Warmth Monitor is proprioceptive because it allows to easily obtaining metric data without any interpretation by the researcher. The Warmth Monitor, or other Feelings Monitors, have been used to measure emotional responses to advertising in a large number of empirical studies in marketing (e.g. Aaker, Stayman and Hagerty, 1986; Stayman and Aaker, 1993; Vanden Abeele and Maclachlan, 1994; Baumgartner, Sujan and Padgett, 1997; Derbais and Poncin, 2005).

3.0 METHODOLOGY

The experiment took place at l'Université du Québec à Rimouski, using 72 night time adult undergraduate students in business administration. Therefore, the data analysis and interpretation of the results originated from two samples of 36 students.

The same three commercials were shown to all the members of both experimental groups. The first one was exclusively used for training purposes, so the participants would be totally at ease working with the Warmth Monitor. This exercise was done three times and after every practice run there was a discussion with the students regarding the recording process, this to ensure that they totally understood the task asked of them. Both groups were then randomly split in two sub-groups, so that the first half of the students would be shown the two commercials in an opposite order to the other half, this to eliminate any warmth-sequence effects across the two commercials (e.g. Vanden Abeele and Maclachlan, 1994; Swaminathan and Kent, 2013).

Also, since it has clearly been established in the literature that the level of viewer involvement with a television program has an impact on advertising effect (e.g. Soldow and Principe, 1981; Krugman, 1983; Norris and Colman, 1993; Tavassoli, Shultz and Fitzsimons, 1995; Owolabi, 2009; Moorman, Willemsen, Neijens and Smit, 2012), it was decided to show the commercials without inserting them into a television program.

4.0 FINDINGS

4.01 PARTICIPANTS

Data on descriptive variables computed for the 72 participants are shown in Table 1.

Table 1: Characteristics of Respondents (n = 72)

	Frequency	%
Gender		
Male	28	38.9
Female	44	61.1
	Average	**SD**
Age	23.3	2.6

The subjects' age varies from 19 to 31 (A = 23.30; SD = 2.60). The sample consisted mainly of women (44; 61.1 %). No significant statistical differences between data coming from the experimental and control groups for both these descriptive variables were found, leading us to conclude that the samples employed for the purposes of this study were similar regarding the age and sex of the participants.

4.01 RELIABILITY OF WARMTH MEASUREMENT

The data coming from both experimental groups were split in half so two parallel series for the mean warmth scores could be computed. Next, for both for the "paper and pencil" and computerised procedures, these series were correlated and split-half reliabilities were calculated. These results are shown in Table 2.

Table 2: Split-half reliability for Warmth

Measure	n	Correlation between halves	Split-half (r)
Warmth (PP)	36	.87	.91
Warmth (C)	36	.91	.95

Note: The measures have "PP" and "C" in parentheses to identify the method used for measuring warmth, respectively the "paper and pencil" and computerised procedures.

On the basis of the results shown in Table 2 it is obvious that, for the sample size used, the warmth measures have sufficient reliability.

4.02 AVERAGE WARMTH SCORES OBTAINED FROM BOTH WARMTH MONITOR PROCEDURES

The results produced by both Warmth Monitor procedures taking the form of vertical lines, selecting the unit of measurement required some choices to be made. In their experiment, (Aaker, Stayman and Hagerty, 1986) measured warmth levels at equidistant points of the vertical axis of the lines drawn for each commercial tested. Using such a procedure can be considered imprecise because the measures obtained in this manner do not represent the same time intervals for every participants or commercial.

To avoid such potential inaccuracies, it was instead decided to divide the vertical axis of the lines into segments that corresponded to five-second intervals. This method produced six warmth level measurements for both 30 seconds TV commercials used in the experiment. The Warmth Monitor's horizontal axis was divided into 10 equally spaced intervals, the "Absence of warmth" given a value of 0 during "Emotional (moist eyes)" was worth 10 points. The results, for both the "paper and pencil" and computerised procedures, can be found in Table 3.

We also checked that both samples were normally distributed for the 24 data series generated in this experiment using the Shapiro-Wilk test, which is well suited to samples of less than 5000 observations. As all the computed p-values were greater than the significance level (i.e. alpha=0.05), one cannot reject the null hypothesis that all the variables from which the two samples were extracted follow a Normal distribution.

Table 3: Average warmth levels and variances from "Paper and Pencil" and computerised Warmth Monitor procedures

	Average warmth scores (PP)		Average warmth scores (C)	
Segment	Commercial A	Commercial B	Commercial A	Commercial B
1	5.23	5.21	5.25	5.26
2	6.18	5.42	6.07	5.39
3	7.32	6.36	7.41	6.38
4	8.48	6.60	8.52	6.57
5	8.25	6.67	8.19	6.70
6	6.76	5.47	6.73	5.42
	Variances (PP)		**Variances (C)**	
Segment	Commercial A	Commercial B	Commercial A	Commercial B
1	0.020	0.110	0.011	0.028
2	0.064	0.389	0.010	0.066
3	0.183	0.887	0.025	0.072
4	0.299	0.900	0.020	0.121
5	0.348	1.061	0.018	0.157
6	0.245	0.898	0.013	0.178

To verify the first hypothesis (H1); that the average warmth scores obtained from both Warmth Monitor procedures are not significantly different, 12 t statistics that follow a Student distribution with ($n_{pp} + n_c - 2$) degrees of freedom were computed. These statistics compare each average warmth level score obtained using the "paper and pencil" procedure with the corresponding score given by the computerised method. Using Student's t test requires a decision to be taken beforehand on whether variances of the samples are to be considered equal or not. Fisher's F tests to verify the hypothesis of equality of the variances were done, and the results were used in the subsequent calculations of the t statistics. If the computed p-value for a Student's t test is greater than the significance level (i.e. alpha=0.05), the null hypothesis (i.e. H_o: The difference between the means is equal to 0) cannot be rejected. The computed p-values, and the corresponding percentage of risk of rejecting the null hypothesis H_o while it is true, can be found in Table 4.

Table 4: Computed p-value for a student's t test and percentage of risk to reject the null hypothesis H_o while it is true for the average warmth scores obtained from both Warmth Monitor procedures.

Segment	Commercial A	Commercial B
1	0.501 (50.14%)	0.635 (63.49%)
2	0.021 (2.09%)	0.601 (60.12%)
3	0.248 (24.84%)	0.810 (81.01%)
4	0.678 (67.79%)	0.861 (86.13%)
5	0.561 (56.09%)	0.915 (91.53%)
6	0.729 (72.88%)	0.777 (77.68%)

Only one of the computed p-values found in Table 4 (i.e. Segment 2 – Commercial A) is inferior to the significance level (i.e. alpha=0.05). For all other cases, none of the results entitles one to reject the hypothesis that the differences of the average warmth scores obtained from both Warmth Monitor procedures are not equal to 0. Also, on 10 of 12 instances, the percentages of risk for rejecting the null hypothesis while it is true are greater than 50%. It cannot be deduced from this outcome that both procedures generated identical warmth levels measurement, but it is reassuring that almost all the observed differences between the two groups are slanted in the right direction.

To verify the second hypothesis (H2); that the measures obtained using the computerised version of the Warmth Monitor vary less between the participants than for those using the "paper and pencil" recording method, Fisher's F tests were computed to establish the equality of the variances between each of the twelve data series obtained using the "paper and pencil" procedure with the corresponding data series given by the computerised method. The p-values (Two-tailed) and the corresponding percentages of risk of rejecting the null hypothesis H0 while it is true can be found in Table 5.

Table 5: Two-tailed p-value for a Fisher's F test and percentage of risk to reject the null hypothesis H_0 while it is true for warmth scores obtained from both Warmth Monitor procedures.

Segment	Commercial A	Commercial B
1	0.062 (6.21%)	0.345 (34.52%)
2	0.018 (1.82%)	0.047 (4.72%)
3	0.016 (1.55%)	0.025 (2.45%)
4	0.048 (4.81%)	0.029 (2.90%)
5	0.000 (0.03%)	0.029 (2.86%)
6	0.011 (1.10%)	0.034 (3.41%)

The data shown in Table 5 reveal that only the results for the first segment of both commercial tested do not confirm that the measures obtained from the computerised version of the Warmth Monitor vary less than those of the "paper and pencil" procedure. This can be explained by the fact that all the participants begin using their version of the Warmth Monitor at the neutral point of the scale, and that the first segment of the commercial only lasts 5 seconds. Hence, it is not overly surprising that no significant differences between the two procedures can be observed for the first segment of the commercials. Nonetheless, the observed difference for the first ad segment between the two groups is fairly important (but less than the significance level of alpha = 0.05). For all the other cases (i.e. segments 2 to 6), the two groups behaved differently in a significant way. Based on this, we consider that it can be asserted that the variance of the scores compiled with the computerised procedure is significantly inferior to the variance of those coming from the "paper and pencil" method.

5.0 CONCLUSION

The results of this study support the hypothesis that are using a computerised version of the Warmth Monitor will produce results that vary less between the participants than those coming from the traditional "paper and pencil" recording method. To some extent they also show that both procedures will generate similar warmth levels measurements.

Several researchers (e.g. Aaker, Stayman and Hagerty, 1986; Vanden Abeele, and Maclachlan, 1994; Derbais and Poncin, 2005) have already listed the many advantages of using process-tracing procedures

over "after the fact" verbal self-reports methods, or physiological measurements. Using the Warmth Monitor method, or another advertising emotional response tracing technique, is very appealing, particularly because of its simplicity and the fact it does not require using complicated and expensive electronic equipment. Although considerable work remains to be done on demonstrating the validity and reliability of the Warmth Monitor, we feel that our research has shown that such a measurement method can be an effective option to consider for decision-making purposes and research.

REFERENCES

Aaker, D. A., Stayman D. M., and M. R. Hagerty (1986). Warmth in advertising: Measurement, impact and sequence effect. *Journal of Consumer Research*, 12, 365-381.

Batra, R., and M. L. Ray (1986). Affective responses mediating acceptance of advertising. *Journal of Consumer Research,* 13, 234-249.

Baumgartner H., M. Sujan and D. Padgett (1997). Patterns of affective reactions to advertisements: the integration of moment-to-moment responses into overall judgments. *Journal of Marketing Research*, 34 (May), 219-232.

Coulter, K. S., (1998). The effects of affective responses to media context on advertising evaluations. *Journal of Advertising*, 27 (4), 41-51.

Derbaix, C. and I. Poncin (2005). La mesure des réactions affectives en marketing : évaluation des principaux outils. Rechercheet Applications en Marketing, 20, 55-75.

Edell J.A. and M.C. Burke (1987). The power of feelings in understanding advertising effects. *Journal of Consumer Research*, 14, 421-433.

Fenwick I. and M.L. Rice (1991). Reliability of continuous measurement copy-testing methods. *Journal of Advertising Research*, February/March, 23-29

Krugman, H. E., (1983). Television Program Interest and Commercial Interruption. *Journal of Advertising Research*, 23 (1), 21-23.

Lioret Segura, S. and Gonzalez-Roma, V. (2003). How do respondents construe ambiguous response formats of affect items? Journal of Personality and Social Psychology, 85 (5), 956-968.

Medianu, S., (2007). Mood Incongruent Recall: An Examination of Mediating Variables. Lethbridge Undergraduate Research Journal, 2, 1-5.

Moorman, M., Neijens, P. C., and E. G. Smit (2012). Program-Involvement Effects on Commercial Attention and Recall of Successive and Embedded Advertising. *Journal of Advertising*, 41(2), 25-38.

Norris, C. E., and A. M. Colman, (1993). Context effects on memory for television advertisements. *Social Behavior and Personality*, 21(4), 279-296.

Owolabi, A. B., (2009). Effect of Consumers Mood on Advertising Effectiveness. Europe's *Journal of Psychology*, 4, 118-127.

Plutchik, R. (1980). Emotion: A Psychoevolutionary Synthesis, Harper and Row, New York (1980)

Richins, M. L. (1997). Measuring emotions in the consumption experience. *Journal of Consumer Research*, 24 (September), 127-146.

Roy, M. (2013). Effects of Commercial Breaks Placement on Television Advertisement Effectiveness. *International Journal of Business and Social Science*, 4(9), 73-79.

Soldow, G. F., and V. Principe, (1981). Response to Commercials as a Function of Program Context. *Journal of Advertising Research*, 21 (2), 59-65.

Stayman, D. M., and D. A. Aaker (1993). Continuous measurement of self-report of emotional response. *Psychology & Marketing*, 10, 199-214.

Swaminathan, S., and R. Kent, (2013). Second-by-second analysis of advertising exposure in TV pods: The dynamics of position, length, and timing. *Journal of Advertising Research*, 53(1), 91-100.

Tavassoli, N. T., Shultz, C. J., and G. J. Fitzsimons, (1995). Program involvement: Are moderate levels best for ad memory and attitude toward the Ad? *Journal of Advertising Research*, 35 (5), 61-72.

VandenAbeele P. and D.L. MacLachlan (1994). Process tracing of emotional responses to TV Ads: Revisiting the warmth monitor, *Journal of Consumer Research*, 20 (March), 586-600.