Does Humor in Radio Advertising Affect Recognition of Novel Product Brand Names?

ERON M. BERG LOUIS G. LIPPMAN

Department of Psychology Western Washington University

ABSTRACT. The authors proposed that item selection during shopping is based on brand name recognition rather than recall. College students rated advertisements and news stories of a simulated radio program for level of amusement (orienting activity) before participating in a surprise recognition test. Humor level of the advertisements was varied systematically, and content was controlled. According to signal detection analysis, humor did not affect the strength of recognition memory for brand names (nonsense units). However, brand names and product types were significantly more likely to be associated when appearing in humorous advertisements than in nonhumorous advertisements. The results are compared with prior findings concerning humor and recall.

Key words: advertising, humor, memory, recognition

RARELY can a person tune in to a radio station without hearing an advertisement that is meant to be humorous. In fact, humor is used in 30.6% of all radio advertising (Weinberger & Campbell, 1991). Although these advertisements may be interesting or even amusing to the listener, do they actually help the advertisers achieve one of their goals—that of product brand recognition? There are inconsistencies in findings that appear to depend on research design, type of humor presented, and type of outcomes measured.

Weinberger and Campbell (1991) evaluated the use and impact of humor in radio advertising using a system "which cross-classifies product decision-making situations along two dimensions, one being high/low involvement and the other thinking/feeling" (p. 45). The involvement dimensions refer to the level of involvement the purchaser must engage in while selecting the product. An example of a

The audiotapes were produced by Jamie Shea (formerly of KVOS TV-12) with support from Western Washington University's Bureau of Faculty Research.

high-involvement, thinking product might be an automobile or other important durable item; examples of low-involvement, thinking items might be gasoline, vitamins, or cold medicine. High-involvement, feeling items include clothes and hair coloring; low-involvement, feeling items include beer and snack food. Weinberger and Campbell claimed that type of humor and type of product are the keys to understanding the role of humor in advertising and asserted that because product and humor types had not been taken into account by earlier studies, results left "a trail of contradictory findings" (p. 44).

Using categorization and analysis of a proprietary database of 2,500 radio advertisements covering more than 35 product categories, Weinberger and Campbell (1991) demonstrated that humor aided recall of low-involvement products. There were significant variations in the incidence of humor for different product types, with 10% of the advertisements classified as humorous for high-involvement, feeling products compared with 39.6% for low-involvement, feeling products. Advertisements having related humor-that is, funny material that is tied to the type of product or brand name-were rated as more persuasive and were recalled more often than advertisements having no humor or unrelated humor. However, advertisements containing unrelated humor were not recalled more than advertisements with no humor or related humor. Weinberger and Campbell indicated that related humor is most effective with low-involvement, feeling products, which is in keeping with humor's relative incidence. It should be noted that Weinberger and Campbell used real radio advertisements. Results of their investigation indicated that advertisements using humor related to the product were more persuasive and were recalled more than advertisements containing unrelated or no humor. Although their data were collected in an ecologically valid manner, their advertisements could not be controlled for length, content, product brand name familiarity, or type of humor presentation. Methodological differences may be responsible for differences in results between their study and earlier studies (Cantor & Venus, 1980; Duncan & Nelson, 1985; Madden, 1982).

Madden (1982) presented a humorous or nonhumorous radio program with either a familiar product or a new product (e.g., sweet acidophilus milk featured in an advertisement having humor related to the product, humor unrelated to the product, or no humor). Thus, for example, a particular audiotape could contain "The Abbott and Costello Show" (as the humorous program) and a humorous ad related to milk (familiar product). Participants were asked to report the ad's effect on their attention, comprehension, retention, and persuasion. Madden presented results showing that advertisements using related humor caught the attention of participants better than unrelated humor or no humor. There was only a slight, nonsignificant trend showing that nonhumorous advertisements were recalled more often than humorous advertisements. Following presentation of each of the commercials, the producer of the commercial was identified (e.g., "This message was brought to you by the California Dairy Association"; Madden, 1982, p. 107). Humor did not affect participants' recall of sources of

advertisements. In addition, recall scores for advertisements presented with humorous and nonhumorous radio shows did not differ. Thus, the lack of humor effect was similar to Cantor and Venus's (1980) evidence that advertisements presented in a serious context were better recalled than advertisements presented in a humorous context.

Cantor and Venus (1980) attempted to make the setting naturalistic in their laboratory study. While in a waiting room, participants were exposed to an audio recording of a fictitious radio broadcast. The broadcast included different types of advertisements (humorous or serious) as well as different radio programs (humorous or serious). Cantor and Venus asked participants to respond to a variety of questions about the broadcast. Results indicated that although humorous advertisements may have led to greater attention than did serious advertisements, nonhumorous advertisements were recalled significantly more often than humorous ones. Furthermore, recall was greatest when the serious ad was presented with the serious radio show. Although the Cantor and Venus study and the Madden study both showed enhanced attention for humorous material, they showed different effects on recall. Madden found no difference, whereas Cantor and Venus found that the nonhumorous material was recalled better.

Duncan and Nelson (1985) also presented a fabricated radio broadcast in a laboratory setting in order to study the effect of humor on attention, communication (recall), persuasion, and antagonization. Using two versions (humorous and serious) of a 60-s radio advertisement embedded in 15 min of music and programming, Duncan and Nelson found that participants attended to and appreciated the advertisement more if it was humorous. They also indicated that participants were less likely to be irritated by the ad if it was humorous, but they found no differences for recall based on the presence or absence of humor. Duncan and Nelson's results are mostly consistent with those of Madden (1982) and Cantor and Venus (1980).

Prior researchers used recall as a measure of memory, with the exception of Berg and Lippman (1997), who used recognition and asserted that it is a suitable measure of memory for at least two reasons. First, if recall and recognition measure different levels of memory, then recall may be an overly stringent measure that could underestimate memory strength. Second, if recall and recognition measure different types of memory, then it would be inappropriate to use recall if recognition is the process by which people identify products while shopping. Shopping in these cases is assumed to be the process by which people identify and select products in stores. People who go to a store seeking a particular product type (e.g., chips) are either browsing aisles in a less focused fashion or are seeking a particular product having a specific brand name. If looking for chips, for example, a customer would go to the chip aisle and as a result of recognizing a familiar brand name may then be more likely to purchase that brand name over an unfamiliar one (Berg & Lippman, 1997). The same process would hold true for the customer who happens to browse in the chip aisle. For either of these

instances, recognition would be the appropriate measure of memory for predicting the likelihood of selection of a product brand for purchase. For recall to be the measure that most appropriately captures the process of shopping and product choice, shoppers would remember the brand name without being cued or reminded. They would then proceed to the store with the intent of finding that specific brand of product.

In an effort to control for a variety of variables, Berg and Lippman (1997) presented participants with simulated radio advertisements for common product types having novel brand names. Each humorous ad (a "one-liner") had its non-humorous counterpart in which the main idea or central theme was retained. For purposes of control, humor was always unrelated to product type or brand name. Individuals provided ratings of perceived humor and then were given a surprise test of recognition for product brand names. A recognition measure was used in this case to mimic more closely the process by which shoppers function in a grocery store. Results indicated that brand names presented with humorous statements were recognized significantly more often than those presented with non-humorous statements. These findings are consistent with the results of Weinberger and Campbell (1991), who found that humor enhanced memory, but not with those of Madden (1982), Cantor and Venus (1980), and Duncan and Nelson (1985), who found that humor had no effect on memory.

Although people in the advertising business deal with campaigns in which the purpose is development of brand name loyalty, the present study was designed to examine the effect of a single presentation of the brand name. It could be asserted that any effect in a single presentation would indicate the nature of effects that could be found in a full campaign. The purpose of the present research was to assess brand name recognition as a function of humor in advertisements. Materials were designed to simulate the way brands are advertised in a radio broadcast in the real world (as opposed to a simple listing of material, as might be done in a typical laboratory study). In addition, material that had been created and calibrated to be humorous and nonhumorous (Schmidt, 1994) was used, and novel and unfamiliar brand names, having minimal associations with material from participants' prior experience, were used.

It was predicted that participants would recognize product brand names that had been presented with humorous advertisements (one-liners) more often than brand names presented with nonhumorous advertisements in the broadcast simulation. Because advertisers may want people to associate their brand names with specific product types, that issue was also considered. In keeping with the prediction of brand name recognition, it was predicted that associations between product brand and type would be stronger for those brands and product types that were presented with humorous material than those presented with nonhumorous material. The chance of finding a humor effect on brand name recognition was expected to be reduced by two major factors: First, unrelated advertisements (Weinberger & Campbell, 1991) were used, and, second, potential interference

could arise from interruption of presentation of advertisements by the three news stories, as takes place in a genuine broadcast.

Method

Participants

Participants in the main study were 60 undergraduate psychology students who volunteered to participate in exchange for additional credit in their courses. Nineteen additional participants provided ratings of news material.

Materials

Professional-quality audiotapes were produced; each was approximately 35 min in duration and included a series of news stories and advertisements. Each advertisement consisted of a brand name, a humorous or nonhumorous one-liner, and a product type.

There were three news stories contained on each tape. They were selected from a pool of eight news stories from National Public Radio's "All Things Considered" show and were as close to neutral as possible for humor, complexity, comprehension, arousal, interest, and memorability, according to ratings by 19 participants. These commentaries' durations ranged from 5 min 30 s to 6 min 40 s.

The fictitious product brand names used in the simulated advertisements were low-meaningfulness paralogs (Noble, 1952) such as gojey, latuk, and flot-sam, which were used to minimize prior associations. Product types were taken from grocery store directories and were limited to common single words while maximizing variety (e.g., cereal, cheese, batteries). Product types were in the low-involvement and feeling categories (Weinberger & Campbell, 1991).

One-liners were taken from Schmidt (1994), who selected humorous statements from original sources and created nonhumorous counterparts for each that retained the content but lost the humor. Those one-liners using negative humor or humor that made fun of specific groups were eliminated, resulting in the selection of 12 of the original 20 for use in the present study.

Brand names (paralogs) were embedded in statements that were paired with the humorous or nonhumorous material to make the advertisements that were approximately 10 s in duration. A cluster of four advertisements was presented between the three commentaries for a total of 12 advertisements.

Before the start of the simulated radio broadcast, there were approximately 30 s of recorded instructions informing participants to listen to the broadcast, pretend that it was real, and rate each segment for its level of amusement. At the start of the simulated broadcast, a fictitious radio station call ("this is KWPS radio") was immediately followed by a segment number ("segment number 1") as a cue for the participant to follow along on a mark-sense answer sheet (Scantron). A

series of four advertisements was located between each of the news stories. Five s of white noise followed each segment, during which participants made their ratings. All advertisements were presented in a female voice.

For recognition testing, the original 12 brand names were intermixed with 12 paralogs not previously presented. Participants were asked to indicate which of the brand names had been presented in the advertisements. Brand names were presented auditorily in the broadcast and visually for the recognition test to more closely simulate the process by which people transfer what they have heard in radio advertisements to the products they see on grocery-store shelves. For the measure of affect toward brand names, participants were presented with 24 brand names (12 old and 12 new) auditorily at a 1-s rate. They were asked to rate their level of "liking" of each brand name. Each brand name was preceded by a number (recorded in a male voice) that allowed participants to follow along with their Scantron sheets. The true/false test for association consisted of 24 statements such as "Gojey is bread; true or false." Each brand name occurred twice, once with the correct product type and once with another product type (e.g., "Gojey is soap").

Rotations

Rotations for the three versions of the radio broadcast were designed such that a particular one-liner was not presented with a particular brand name or product type more than once in the study. Even though the lowest meaningfulness paralogs had been selected for use, they still differed in rated meaningfulness. Thus, careful rotation of material was important. Within each combination, product types, brand names, and one-liners were rotated in order of presentation for each of the three audiotapes. To arrange three commentaries in three rotations, a Latin square counterbalancing technique (Keppel, 1991) was used.

Design and Procedure

A within-subject design was used, which has been shown to be more sensitive than between-subjects designs to detect humor effects (Schmidt, 1994). All testing was done in groups ranging from 2 to 29 participants. The sequence of activities was rating each segment for its level of amusement (to ensure exposure to the material), a surprise recognition test, ratings of affect toward the brand names, and finally the test of association between product brands and types.

Participants were given aural instructions on the tape, a sample of the material, and a Scantron. Participants were asked to listen and to pretend that the recording was an actual radio broadcast, and to follow along so they could make amusement ratings for each segment on their Scantron forms. Additional recorded instructions were given before each of the three other tasks. Participants were then presented with one of the three versions of the radio broadcast. While listening to the broadcast, participants rated the commentaries and advertisements

for "level of amusement," as an orienting activity. Ratings were done using a 5-point Likert-type scale ranging from *not amusing* (1) to *very amusing* (5).

Once participants had listened to the radio broadcast, they were given a list of the 24 paralogs (12 not previously presented and 12 that had appeared as product brand names in the advertisements) face down on their desks for the surprise recognition test. After a delay of approximately 30 s following the broadcast, participants were instructed to turn over the page and indicate, by darkening the corresponding oval that represents "Y" for yes or "N" for no on the Scantron, which of the 24 "brand names" they had heard before.

After the surprise test for recognition of product names, participants were asked to rate their affect toward the brand names by indicating how much they "liked" each brand name. Finally, the true/false test to assess strength of association between product types and their brand names was administered; in it, participants decided which pairs had actually been presented in the radio broadcast.

Results and Discussion

Amusement

Reactions to material known to be humorous and nonhumorous (Schmidt, 1994) were compared. Amusement ratings of sentences that Schmidt had identified as humorous or not humorous were averaged separately, yielding M = 1.89, SD = 0.63 for humorous items and M = 2.06, SD = 0.70 for nonhumorous statements. Nonhumorous material was rated as significantly more amusing than humorous material, F(1, 59) = 5.80, p = .019, MSE = 0.14, indicating that participants evidently considered "amusement" to be something different from "humor." It should be noted, however, that participants produced amusement ratings for all segments-advertisements as well as news stories. The more lengthy and engaging news stories could certainly have affected perception of the one-liners. In the context of a simulated radio broadcast, perhaps a neutral statement seemed unusual or novel, inducing an interpretation of "amusing." Amusement ratings of the advertisements ranged only from 1 to 3, with 95% of participants responding with either 1 or 2, indicating that participants found little entertainment with the material. A point-biserial correlation between Schmidt's categories (humorous or not humorous) and participants' ratings of amusement revealed no relationship between the variables. It is possible that participants based their ratings on "strangeness" or "novelty" with some of the advertisements that included nonhumorous sentences embedded in advertisement text. It is also possible that the sentences functioned differently as a result of being embedded in advertisements.

Recognition

Participants' decisions were subjected to signal detection analysis. Briefly, in the present instance, a hit rate was correct identification of a paralog as

having appeared in one of the advertisements during the simulated radio broadcast and a false alarm was incorrect judgment that a new paralog had been one of the brand names in one of the radio advertisements. Hit rate, false alarm rate, and d' were each computed overall as well as for humorous and nonhumorous material.

The overall hit rate was M = .41, SD = .17, the false alarm rate was M = .39, SD = .16, and average d' = .095 (SD = .60). The hit rate for humorous material (M = .40, SD = .20) was not significantly different from the hit rate for nonhumorous material (M = .43, SD = .21), F(1, 59) = 0.69, p = .41, MSE = 0.03. The d' for humorous material (M = .112, SD = .65) was not significantly different from the d' for nonhumorous material (M = .172, SD = .74), t(1, 59) = -0.47, p = .64, MSE = 0.48, or r = .04, indicating no relationship between humor and memory. Because participants' amusement ratings did not correspond with humor based on prior research, correlations were computed between participants' hit rates and their ratings of amusement for each advertisement. Correlations ranged from r = -.22 to r = .44 with a median r = .07, indicating that there was no relationship between amusement and brand name memory.

The primary prediction for this study was not supported by the results, in that participants did not recognize brand names associated with humorous material any more often than they recognized brand names associated with nonhumorous material. Several explanations for these results are consistent with the literature. First, a memory effect for humorous material might not be detectable due to the difficulty of the task; material to be remembered was unfamiliar and embedded in "strange" advertisements. Furthermore, the advertisements were interrupted by three news stories that were each 5 min in duration. Second, these results are consistent with those of Madden (1982) and Duncan and Nelson (1985), who found that the presence or absence of humor did not affect memory. Third, it is possible that the humorous material identified by Schmidt (1994) may have operated differently in the context of a radio advertisement. Finally, it is also possible that the shopping model was not adequately tested in this study because participants were not asked directly about their decision-making processes, but only about their ability to differentiate visually between brand names of which they had or had not heard.

Liking

Ratings of participants' "liking" of brand names were analyzed twice. First, they were divided into ratings of "new" brand names (those not presented in the radio show) and ratings of "old" brand names (those that had been presented in the broadcast). Ratings of new brand names (M = 2.15, SD = 0.69) differed significantly from ratings of old brand names (M = 2.01, SD = 0.61), F(1, 59) = 7.24, P = .009, MSE = 0.09. This outcome indicates that participants liked brand names that had not been presented in the radio show more than those that

had been. It is possible that participants were responding to new material more positively because it was novel and they were satiated by the paralogs that had been presented twice during the simulated broadcast. Although nonsense words lose their novelty very quickly, when first presented they often provoke interest, curiosity, and even some amusement. The relatively small difference in means (2.15 for new brand names and 2.01 for old brand names) as well as the low mean values also indicate that participants did not express positive affect about product brand names.

Brand names presented with humorous advertisements were separated from those presented with nonhumorous advertisements and their ratings were compared. Ratings of brand names that appeared with humorous material (M = 1.94, SD = 0.64) did not differ significantly from ratings of brand names that appeared with nonhumorous material (M = 2.08, SD = 0.72), F(1, 59) = 3.02, p = .087, MSE = 0.18. Thus, affect toward the brand name was not affected by the humor level of accompanying material. It is important to note that the absence of an affective outcome may be due to the fact that the humor was not related to the product. Note that Weinberger and Campbell (1991) found an affective difference only when the humor used was related to the product.

Association

The true/false test of association was designed such that a signal detection analysis could be conducted. Of the 24 true/false statements, 12 were true and 12 were false. However, only 12 product brand names were presented. Each brand name was presented with two product types: one with which it had appeared and one with which it had not appeared. This allowed an overall hit rate, false alarm rate, and d' to be computed as well as a hit rate, false alarm rate, and d' for those brand names that had appeared with humorous material and those that had appeared with nonhumorous material.

The overall average hit rate was .42 (SD = .16), the false alarm rate was .40 (SD = .18), and d' was .04 (SD = .72). This overall d' indicates weak memory strength for this association task. However, the hit rate for humorous material (M = .49, SD = .20) differed significantly from the hit rate for nonhumorous material (M = .34, SD = .18), F(1, 59) = 40.52, p < .001, MSE = .02. The false alarm rate for humorous material (M = .38, SD = .22) did not differ significantly from the false alarm rate for nonhumorous material (M = .42, SD = .21), F(1, 59) = 1.45, P = .233, MSE = 0.03. The d' for humorous material (M = .29, SD = .69) differed significantly from the d' for nonhumorous material (M = .17, SD = .58), t(1, 59) = 3.92, p < .001, MSE = 0.41. Participants were able to detect correct associations between product brand and type more often when those brand names had been advertised with humorous material than when they had been advertised with nonhumorous material.

Conclusions

Amusement and Liking

Amusement results indicated that participants rated advertisements that were intended to be nonhumorous as more amusing than those advertisements intended to be humorous. Although participants may have interpreted amusement as something other than humor in the context of these advertisements, it seems likely that judgments of news stories along with advertisements may have overshadowed and distorted the ratings of amusement.

Results of the liking data indicated only that participants liked brand names that had not been in the advertisements more than brand names presented in the advertisements, which seems consistent with participant boredom or responsivity to novel items. There were no differences in level of liking for brand names presented with humorous material versus brand names presented with nonhumorous material, which is consistent with Weinberger and Campbell's (1991) finding of an affective difference for products advertised with related humor only. Unrelated humor was a constant condition in the present study. Advertisers would be well advised to consider the humor (related or unrelated to brand name) used in their advertisements, as some advertisements have the ability to detract from participants' affect toward brand names.

Memory: Recognition and Association

The prediction that product brand names associated with humorous material would be recognized more often than product brand names associated with nonhumorous material was not supported. Because Madden (1982) and Weinberger and Gulas (1992) found that unrelated humor was much less effective than related humor for enhancing memory, the present study was designed to provide conservative estimates of humor effects.

The second prediction that associations between product brand names and product types would be stronger or more likely when the material was presented in humorous context than nonhumorous context was supported. Participants showed more associations between product brand names and product types when the brand name and type were presented with humorous material rather than non-humorous material.

Differentiation of correct product brand name/product type combinations from incorrect pairings is an indication of enhanced memory as a result of the advertisement. Thus, even if results of the recognition test did not provide evidence of a humor effect on brand name recognition, the test of association at least provides evidence of a humor effect. It is of particular interest that this advantage arose from advertisements in which the humor was not connected directly to the product.

Shopping Model

According to the shopping model on which this study was based, people use recognition memory rather than recall of brand names to determine product selection. In other words, people look for a product type and then recognize a brand name rather than recall a specific brand name and seek it for purchase. To test this shopping model, we assessed recognition memory for product brand names. As has been noted, overall strength of memory for brand names was low. For the shopping model to operate as it was suggested, strength of memory would need to be substantially higher. And it probably is in the real world, where advertisements are far more lengthy and brand names are repeated in advertising campaigns. In addition, participants were generating ratings rather than participating in a contrived shopping situation. Thus, it is very possible that the present test was inadequate. A more ideal evaluation would take place in a more naturalistic setting-for example, actual advertisements presented without the participant knowing it was a research project, as done by Cantor and Venus (1980), and an observational design (in the grocery store) with a questionnaire after the products had been purchased to determine the process of selection/memory.

Final Comments

The present study was neither a "pure" lab test, in which participants read a list of contrived advertisements and then were tested, nor was it an "ecologically valid" study that allowed participants' familiarity with advertisements and brand names to be confounding factors. Rather, participants listened to a fictitious radio broadcast, and memory was evaluated by a surprise recognition test. As a constant condition, product type, content of the humorous messages, and novel brand names were all unrelated. These conservative tactics provided a nice example of a well-controlled simulation. The recognition and association tasks both required participants to transfer auditory information from the advertisement to the tasks that were presented in the visual mode. It had been reasoned that this translation process has to underlie the way radio advertising functions. All of these elements indicate that any humor effect found would be an underestimate of the effect realized in an actual advertising campaign.

REFERENCES

Berg, E. M., & Lippman, L. G. (1997, April). The effect of humor in radio advertising on novel product brand recognition. Poster session presented at the annual meeting of the Western Psychological Association, Seattle, WA.

Cantor, J., & Venus, P. (1980). The effect of humor on recall of a radio advertisement. Jour-

nal of Broadcasting, 24, 13-22.

Duncan, C. P., & Nelson, J. E. (1985). Effects of humor in a radio advertising experiment. Journal of Advertising, 14, 33-40, 64.

Keppel, G. (1991). Design and analysis: A researcher's handbook (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Madden, T. J. (1982). Humor in advertising: Applications of a hierarchy of effects paradigm. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Noble, C. E. (1952). The analysis of meaning. Psychological Review, 59, 421-430.

Schmidt, S. R. (1994). Effects of humor on sentence memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20, 953–967.

Weinberger, M. G., & Campbell, L. (1991). The use and impact of humor in radio advertising. *Journal of Advertising Research*, 31, 44–52.

Weinberger, M. G., & Gulas, C. S. (1992). The impact of humor in advertising: A review. *Journal of Advertising*, 21, 35–59.

Manuscript received October 1, 1999 Revision accepted for publication October 16, 2000