

DEVELOPING COPY TESTS THAT ESTIMATE BRAND USAGE

Brian Wansink
Assistant Professor of Business Administration
Amos Tuck School -- Dartmouth College

Michael L. Ray
John G. McCoy/Bank One Professor of Innovation and Creativity in Marketing
Graduate School of Business
Stanford University

The article has been published. It's appropriate citation is . . .

Wansink, Brian and Michael L. Ray (1997), "Developing Copy Tests that Estimate Brand Usage," in William Wells (Ed.) Measuring Advertising's Effectiveness, Cambridge, MA: Lexington, 359-370.

Acknowledgements

The authors are grateful to the Stanford Graduate School of Business Doctoral Program, and the Amos Tuck School for the generous financial support they have provided. The first author also gratefully acknowledges financial support received in conjunction with the Marketing Science Institute's Alden G. Clayton Doctoral Dissertation Proposal Competition award.

DEVELOPING COPY TESTS THAT ESTIMATE BRAND USAGE

ABSTRACT

When the objective of an advertising campaign is to increase the consumption or usage frequency of a package good, copy testing measures must be sensitive to this objective. Since measuring actual usage can be prohibitive in terms of time and money, we present two measures which can be examined in a laboratory -- cognitive responses and usage-intentions. Two different laboratory studies underscore the factors that influence the sensitivity and validity of these two different types of measures. With measures of usage intention, one's usage frequency of the target brand determines the appropriate measure question. With cognitive responses, the sensitivity is determined by the specificity of the elicitation procedure. Methodological and measurement implications are suggested along with a general method for further testing of stylized measures such as usage.

The growing interest in brand equity is resulting in new advertising objectives for some brands of consumer packaged goods. Some campaigns emphasize using the brand instead of simply choosing it over a competing brand. That is, many versatile, high-penetration packaged brands are well-suited for advertising campaigns that encourage loyal consumers to use the brand more frequently, or in new ways (Wansink and Ray 1993). This chapter describes the copy-testing measures that most accurately predict whether a campaign will generate increased usage. Such usage-related campaigns are most commonly employed by mature, dominant brands (such as Campbell's Soup); high loyalty niche brands (such as Grey Poupon); and industry associations (such as the American Dairy Council).

Both industry professionals and academicians have criticized traditional copy-testing methods for their inability to capture usage-related responses accurately (Marketing Science Institute 1983). Measures of purchase intention and brand attitude, for instance, are no longer satisfactory when our interest is in usage. These measures are too insensitive when consumers are brand loyal. Brand attitude measures are often at a ceiling. Furthermore, estimates of purchase intentions become unreliable when the product is already in inventory and must be depleted before it can be repurchased. Indeed, recent studies have shown the correlations between measures of brand attitude and subsequent usage to range from -.10 to .23 for heavy users (Wansink and Ray 1992).

This chapter describes research which examines the validity of different types of usage-related measures that can be reasonably collected in a copy-testing environment. From a managerial standpoint, we suggest measures that would

determine which of two campaigns (such as a “Mmm Mmm Good” campaign or a “Soup is Good Food” campaign for Campbell’s Soup) would be more effective at increasing usage. In particular, this paper makes three specific contributions. First, it provides a basic framework for understanding the sequential effects of advertising on usage-related responses. Second, it suggests valid quantitative measures that estimate an ad’s impact on usage. Last, it suggests cognitive-response elicitation questions that are most sensitive to usage.

After outlining a basic framework for usage-related responses, the results of two copy-testing studies will be described. The first study shows the quantitative measures that correlated most highly with usage. The second study shows a method that maximizes the sensitivity of cognitive responses toward usage-related thoughts.

HOW ADVERTISING IMPACTS USAGE

An ad can be processed in either a peripheral way or a central way, depending on how much attention a consumer allocates to the ad (Petty, Cacioppo, and Schumann 1989). Suppose a person spends little time thinking about an ad he or she has just seen. Even if exposure to this ad has no influence on consumer attitudes toward the brand (A_{brand}), the ad may still have an impact on consumption or usage if it simply raises his or her awareness of the brand (Ehrenberg 1974), or increases the chance that it might be considered for usage (Nedungadi 1991). Consistent with this, a field study conducted by Seagrams suggested that changes in attitude need not precede changes in the usage frequency of a product. This study analyzed aggregated data from brand loyal consumers (Schiller, Schribner and Belkin 1982), showing that consumers who were frequently exposed to Seagrams ads also consumed more the their products.

This increase in usage occurred “in the absence of any related increase in product beliefs or in attitude.” (See Figure 1.) Although this study does not prove causality, its results are consistent with the notion that changes in usage may occur without accompanying changes in attitude.

Insert Figure 1 About Here

The more frequently studied route to persuasion is the central route. It suggests that when a consumer views an ad, and is highly involved with it, he or she silently generates thoughts about the brand, and these thoughts either alter or fortify beliefs and attitudes about the brand and its use. In turn, these beliefs and attitudes influence usage intentions which eventually influence usage. This perspective is consistent with learning theory and what we would expect when consumers process informational ads (Rossiter and Percy 1988). Copy-testing measures that examine centrally-processed measures will be the focus of this chapter. The low-involvement processing conditions that exist with peripherally-processed messages will not be further examined in this article.

Figure 1 shows the points at which we could sample the ad’s impact. As researchers, our interest is in determining the measures that suggest future usage, while eliminating the time and expense of having to collect actual usage measures. It is important to realize that two broad types of measures of ad effectiveness can be taken in the laboratory without having to take actual usage measures in the field: (1) cognitive responses, and (2) quantitative measures of usage intentions. Both of these can be collected quickly and inexpensively, and can serve complementary purposes.

DETERMINING THE QUANTITATIVE MEASURES THAT PREDICT USAGE

Measures of one's usage intentions (for a particular time period, such as "within the next two weeks") can be obtained either through likelihood measures, or through estimates of one's usage volume. Likelihood measures can be directly obtained by asking an individual how likely ("Highly Unlikely" = 1 to "Highly Likely" = 7) it will be that he or she uses the brand within an upcoming time period. Usage intentions can also be measured by asking one to estimate the volume of a brand he or she might possibly consume within a similar time period.

These two different measures of usage intent have different relative strengths. With infrequent users of a brand, volume estimates will be skewed toward 0 units (especially over a relatively short period of time). This is partially a drawback of numerical estimates that provide no gradation between 0 and 1 unit. In such cases, volume estimates would provide less variance and less information than an estimate of usage likelihood. As a result, usage likelihood estimates would allow a greater gradation in response and would be more sensitive in detecting any potentially different effects these ads might have on usage.

In contrast, with frequent or heavy users of a brand, a volume estimate is likely to be more accurate than a likelihood estimate. This is because the distribution of these volume estimates is more likely to be normally distributed (Pearl 1981). As a result, a volume estimate of one's usage intent is likely to provide more variance and more information about the intended usage of heavy users than is a likelihood measure, which would undoubtedly be at or near 1.0

(100 percent probable). Under these circumstances, volume estimates would be a more accurate estimate of a heavy user's usage volume of a brand.

Empirical Findings

The effectiveness of these different measures was examined by Wansink and Ray (1992) when they exposed 239 subjects from Parent-Teacher Associations to a series of ads for one of three different brands (Campbell's Soup, Jell-O Brand Gelatin, and Ocean Spray Cranberry Sauce). The correspondence between intentions and usage was most impressive when the subjects were segmented into heavy users and light users based on their prior year's usage of the brand. Consumers who consumed more than the median amount for each brand were classified as relatively "heavy users," and the rest as "light users" (Jacoby and Chestnut 1978).

In general, both measures of usage intention (likelihood and volume estimates) were effective in predicting subsequent usage, depending upon how frequently one has tended to consume the brand in the past. As shown in Table 1, heavy users of the brands were more accurate in estimating their usage volume than in estimating their "likelihood" of using these three products ($r = .62, .46,$ and $.23$). In contrast, light users of the brands were unable to accurately estimate their usage volume but were instead much more accurate in estimating the usage likelihood ($r = .42, .78,$ and $.49$). When contrasted with research that indicates that usage volume predictions are often very low (Pilgrim 1957; Cassidy 1981), the results from Table 1 show that volume predictions can be very accurate when frequent users are examined. The results of this study are described in more detail elsewhere (Wansink and Ray 1992).

Insert Table 1 About Here

Implications for Increasing Predictive Validity

These results make two important illustrations. First, brand attitude measures will not always be sensitive enough to usage-related responses from ads. Second, usage intentions can be measured through likelihood estimates or through volume estimates, and each measure is effective under different circumstances. As seen in Table 1, heavy users of a given brand are most accurate when predicting their future usage volume. Light users are most accurate when predicting their likelihood of consuming the brand.

These results can be extended to entire product categories. That is, if a researcher is trying to estimate the impact that an ad will have on the usage of a product category that, relative to other categories, is infrequently consumed, likelihood measures may be more generally accurate than volume measures. However, if the product category is one that, relative to other categories, is frequently consumed, volume measures may be more accurate. In this study, for example, the typical household ate 29.1 cans of soup per year, but only 2.7 cans of cranberry sauce. Given a larger sample of subjects, we would likely find that soup is a product category where usage intentions are best estimated through volume measures, while usage intentions for cranberry sauce would be best estimated through likelihood measures (see Figure 2). This relationship should be even stronger when examining the heavy users of a frequently consumed category, or when examining the light users of an infrequently consumed category. Indeed, as can be seen in Table 1, volume estimates provided relatively accurate estimates of usage for heavy users ($r = .62$) of soup and likelihood estimates provided relatively accurate estimates of usage for light users ($r = .49$) of cranberry sauce.

Insert Figure 2 About Here

The findings described here underscore the importance of usage-related measures over the simple measures of attitude that are typically collected during copy-testing. Specifically, it is important to understand that volume estimates best approximate the actual usage of heavy users (or of frequently consumed brands) and that likelihood estimates are best used with light users (or with infrequently consumed brands). More specific diagnostic information, such as usage-related thoughts and feelings, can be obtained by examining the specific thoughts that consumers generate when viewing these ads.

ELICITING USAGE-RELATED COGNITIVE RESPONSES

Understanding the effectiveness of an ad is greatly aided by knowing a consumer's thoughts as he or she views it. These thoughts help us better estimate the impact these ads will have on attitudes and usage, and they also suggest ways in which the ads can be changed to be more effective. Unfortunately, the traditional procedure by which these thoughts are elicited may not yield valid or reliable findings (Russo, Johnson, and Stephens 1989).

The initial research with cognitive responses (or verbal protocols) was pioneered by Greenwald (1968) and then introduced into advertising by Wright (1973). Their work indicated that cognitive responses can mirror the actual thoughts that occur to people as they evaluate a persuasive message. In these studies, cognitive responses are typically elicited with instructions such as, "Write down any thoughts that went through your mind while reading the ad." These

written thoughts are typically coded as either counterarguments, support arguments, or source derogations (Smead, Wilcox, and Wilkes 1981; Wright 1973).

One problem with this coding scheme is that it does not specifically address thoughts that are usage-related, nor does it necessarily encourage ones that could be of diagnostic value. Although a multitude of thoughts may be generated as one views an advertisement, only a small percentage of them will actually be communicated (Wright and Rip 1980; Kidder 1980). After subjects see an ad, they are typically asked to record their thoughts when viewing it. These instructions are general, and a portion of the random thoughts that results could be minimized if subjects had a better idea of what is expected of them (Ericsson and Simon 1984). In short, when a researcher is focusing on usage-related thoughts, the conventional procedure of simply asking for general reactions may not be as useful as procedures or questions that are less ambiguous.

Two Options for Eliciting Usage-related Responses

A person viewing an ad may generate many thoughts about cognitive responses, but not all of them will be communicated because of time constraints or cognitive capacity constraints (Ericsson and Simon 1984). To uncover these thoughts about a particular target issue, researchers have used either pre-exposure elicitation exercises, or directed post-exposure instructions.

If a subject is given no instructions prior to their exposure to an ad, he or she is free to think of any issues that come to mind. Pre-exposure elicitation exercises (such as practice tests or examples) frame a subject's processing by suggesting a range of issues which one might consider. One way this can be accomplished is by providing subjects with a hypothetical example or illustration of what another subject might have written when he or she viewed a related ad

(Keller 1987). A second way this is accomplished is by providing subjects with a practice trial that is followed with standardized feedback. The feedback, for instance, can be presented in the form of a pre-written checklist which instructs them to reread their responses to insure they are not simply writing down a replay of the ad (Batra 1984).

Providing subjects with pre-exposure elicitation exercises intensifies their processing of these target issues during exposure. In contrast, giving directed post-exposure instructions to subjects after they view an ad encourages them to cognitively edit their less relevant thoughts before writing them down. One way this can be accomplished is by instructing subjects to address specific issues of interest (Wright 1980). For instance, a researcher can ask subjects how they feel about using the product, if they agree or disagree with the ad, or if it reminds them of any past experiences with the product (Wright and Rip 1980).

Pre-exposure elicitation exercises, and directed post-exposure instructions both share risks of potential reactivity. The primary concern is that these procedures may “force” a subject to generate thoughts about a particular target issue that would have otherwise never occurred to them (Turner 1988; Nisbett and Wilson 1977). As a result, such thoughts would be invalid, and would bias outcome measures such as beliefs, attitudes, or intentions. A direct way of testing for reactivity is by measuring the impact these different procedures have on critical outcome variables (Russo, Johnson, and Stephens 1989). Nonreactive procedures should have no influence the ratings of outcome variables when compared to that of a control group. In other words, if these different procedures are nonreactive, there should be no difference in the ratings of A_{brand} , A_{ad} , and usage intentions between subjects who are given pre-exposure elicitation exercises, directed post-exposure instructions, or neither.

Empirical Findings

The effectiveness of these two different elicitation methods was examined in a study that involved 74 adults who were recruited from Parent-Teacher Associations and who were given \$6.00 for their effort (see Wansink, Ray, and Batra 1994 for details). This study found that using either pre-exposure elicitation exercises or directed post-exposure instructions increased the number of usage-related thoughts generated by subjects, but was not reactive. That is, there were no corresponding differences in the ratings of A_{brand} , A_{ad} , or usage intentions between subjects who were given pre-exposure elicitation exercises, directed post-exposure instructions, or neither.

In a general sense, these results are consistent with what Batra (1984) found when examining different types of elicitation exercises for different dependent variables. Batra's results showed that general instructions can be as effective as directed instructions, but only when accompanied by some form of vivid pre-elicitation exercise or illustration, such as an example or as a practice trial.

When should pre-exposure exercises be used in favor of directed post-exposure instructions? It is important to realize that both options are not always available. Involving subjects in pre-exposure exercises is not always feasible, and it can be constrained by the experimental design or time limitations. Under such circumstances, directed post-exposure instructions are the best alternative. When pre-exposure exercises can be used, Figure 3 suggests that they might elicit more thoughts about a target issue. It is important to note that the combination of the two procedures, however, provides no greater sensitivity than does either by itself.

Insert Figure 3 About Here

A General Method for Increasing Cognitive Response Sensitivity

Research dealing with cognitive responses is important because of the generalizations it makes regarding the cognitive response sensitivity (see also Wansink and Ray 1994) . In doing so, it suggests a general pre-testing method that can help researchers determine what procedure will be most appropriate for eliciting usage-related cognitive responses. The general four step method follows:

1. Select a number of pre-exposure elicitation exercises and directed post-exposure instructions believed to provide the greatest level of sensitivity toward usage-related responses. Be certain to include a control condition.
2. Design the study by having the various procedures under examination represent between-subjects factors. Statistical power can be increased by having subjects respond to multiple ads. Care should be taken to insure that subjects are from a comparable pool as those who will be involved in the future studies.
3. Include outcome variables of interest to insure that the different procedures do not generate reactivity (such as A_{brand} , A_{ad} , usage intentions).
4. Select the elicitation procedure that best achieves the objectives of the study without affecting outcome variables relative to the control condition. For instance, an objective may involve selecting the procedure which maximizes usage-related thoughts, while minimizing unrelated thoughts such as ad playback.

This section emphasizes the importance of increasing the usage-related sensitivity of cognitive response elicitation procedures. Furthermore, it illustrates the steps a researcher must go through if he or she wishes to develop a stylized elicitation procedure for his or her own program of research. The study described here is taken from an ongoing program of research which suggests that either pre-exposure elicitation exercises (such as practice trials or prior exposure), or directed post-exposure instructions can increase this sensitivity without appearing

to be reactive. The combination of the two procedures, however, provides no greater sensitivity than does either by itself.

SUMMARY

When the objective of an advertising campaign is to increase the usage frequency of a packaged good, copy testing measures must be sensitive to this objective. Since measuring actual usage can be prohibitive in terms of time and money, two more primary measures -- cognitive responses and usage-intention measures -- show promise because of their validity and diagnostic value.

These findings underscore the importance of taking usage-related measures in copy tests, instead of simply attitude measures of purchase intentions. Specifically, it is important to understand that volume estimates best approximate the actual usage of heavy users, or frequently consumed brands and that likelihood estimates are more accurate with light users, or with infrequently consumed brands. Additional information about these ads can be obtained by examining the thoughts or cognitive responses that are generated by these ads.

These cognitive responses can best be examined using either pre-exposure elicitation exercises, or directed post-exposure instructions. Early evidence suggests that either procedure can effectively increase the number of thoughts a respondent communicates about a particular target issue, and that they may not be reactive.

Figure Captions

Figure 1 Usage Intention Measures that Correspond Most Closely to Actual Usage

Heavy Users Light Users

Frequently Used Category Infrequently Used Category

Relative Usage Frequency of the Category

Volume Estimates Likelihood Estimates

Use Both Volume & Likelihood Estimates

Figure 2 Types of Cognitive Responses Generated by Various Elicitation Methods

Number of Cognitive Responses

General Instructions Directed Instructions

Example Practice Trial Control Condition

Example Practice Trial Control Condition

Usage-Related Cognitive Responses

Product-Related Cognitive Responses

Irrelevant Cognitive Responses (e.g., ad playback)

References

- Batra, R. (1984). 'Low Involvement' Message Reception -- Processes and Advertising Implications. Unpublished dissertation, Graduate School of Business, Stanford University, Stanford, CA 94305.
- Cassidy, C. M. (1981). Collecting Data on American Food Consumption Patterns: An Anthropological Perspective. In Food and Nutrition Board National Research Council, Assessing Changing Food Consumption Patterns, Washington, D.C.: National Academy Press, 135-154.
- Ericsson, K. A. & Simon, H. A. (1984). Protocol Analysis: Verbal Reports as Data, Cambridge, MA: The MIT Press.
- Ehrenberg, A. S. C. (1974). Repetitive Advertising and the Consumer. Journal of Advertising, 14(2), 25-34.
- Greenwald, A. G. (1968). Cognitive Learning, Cognitive Response to Persuasion, and Attitude Change. In A. G. Greenwald, T. C. Brock, & T. C. Ostrom, (Eds.), Psychological Foundations of Attitudes (pp. 63-102). New York: Academic Press.
- Jacoby, J. & Chestnut, R. W. (1978). Brand Loyalty: Measurement and Management. New York: John Wiley and Sons.
- Keller, K. L. (1987). Memory Factors in Advertising: The Effect of Advertising Retrieval Cues on Brand Evaluations. Journal of Consumer Research, 14 (December), 316-324.
- Kidder, L. H. (1980). Research Methods in Social Relations. New York: Holt, Rinehart and Winston.

- Marketing Science Institute (1983). Beyond Recall. Cambridge, MA: Marketing Science Institute.
- Nedungadi, P. (1990). Recall and Consumer Consideration Sets: Influencing Choice Without Altering Brand Evaluations. Journal of Consumer Research, December, 263-276.
- Nisbett, R. E. & Wilson, T. D. (1977). Telling More Than We Can Know: Verbal Reports on Mental Processes. Psychological Review, 84, 231-259,
- Pearl, R. B. (1981). Possible Alternative Methods for Data Collection on Food Consumption and Expenditures." In Food and Nutrition Board National Research Council, Assessing Changing Food Consumption Patterns, (pp. 198-203). Washington, D.C.: National Academy Press.
- Pilgrim, F. J. (1957). The Components of Food Acceptance and Their Measurement. In Josef Brozek (Ed.), Symposium on Nutrition and Behavior, (pp. 69-73). New York, National Academy Press.
- Rossiter, J. R. & Percy, L. (1987). Advertising and Promotion Management. New York: McGraw-Hill.
- Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The Validity of Verbal Protocols. Memory and Cognition, 17 (6), 759-769.
- Schiller, C., Schribner, R. J., & Belkin, M. (1982). A Study of the Effectiveness of Advertising Frequency in Magazines. Prepared by Time Incorporated in association with Joseph E. Seagram and Sons, Inc. New York: Time Inc.
- Smead, R. J., Wilcox, J. B., & Wilkes, R. E. (1981). How Valid are Product Descriptions and Protocols in Choice Experiments? Journal of Consumer Research, 8 (June), 37-42.
- Stewart, D. W. (1989). Measures, Methods, and Models in Advertising Research. Journal of Advertising Research, June/July, 54-60.

- Turner, C. K. (1986). Don't Blame Memory for People's Faulty Reports on What Influences their Judgments. Personality and Social Psychology Bulletin, 14, 622-629.
- Von Gonten, M. (1989). Personal conversation, August 1.
- Wansink, B., Ray, M. L., & Batra, R. (1994). Increasing the Sensitivity of Cognitive Responses. Journal of Advertising, 24 (2), forthcoming.
- & ----- (1992). Estimating an Advertisement's Impact on One's Consumption of a Brand. Journal of Advertising Research, 32 (May/June), 9-16.
- & ----- (1993). Extension Advertising's Impact on Brand Equity. In D. A. Aaker & A. L. Biel (Eds.), Advertising and Building Strong Brands (pp. 177-194). Cambridge, MA: Lexington Books.
- & ----- (1992). Goal-Related Consumption and Extension Advertising: The Impact on Memory and Consumption. In John Sherry and Brian Sternthal (Eds.) Advances in Consumer Research, 19 (pp. 806-812). Provo, UT: Association for Consumer Research.
- Wright, P. L. (1973). The Cognitive Processes Mediating Acceptance of Advertising. Journal of Marketing Research, 10 (February), 53-62.
- (1980). Message-Evoked Thoughts: Persuasion Research Using Thought Verbalizations. Journal of Consumer Research, 7 (September), 151-175.
- & Peter Rip (1980). Retrospective Reports on Consumer Decision Processes. In J. C. Olson (Ed.), Advances in Consumer Research, 7 (pp. 146-147). Ann Arbor, MI: Association for Consumer Research.

Wansink and Ray found that with both heavy and light users, measures of attitude (both pre-post difference measures and measures of A_{brand}) were very weak predictors of usage. This was not surprising since brand attitude is distantly linked to usage, and is mediated by both the salience of the brand and by one's family's interest in eating it. That is, one may love liver and onions, but would never serve it because no one else in the family likes it (Wansink and Ray 1991).