

# Radio's ROI Advantage



A Major New Study Of  
**Radio's Return On Investment Compared To Television**  
Conducted By Millward Brown and Information Resources, Inc.  
For The Radio Ad Lab  
2004–2005

Radio Ad Lab  
125 West 55<sup>th</sup> Street, 21<sup>st</sup> Floor  
New York, NY 10019

Phone: (800) 364-3239

Info@RadioAdLab.org  
<http://RadioAdLab.org>

**RadioAd**  
Lab

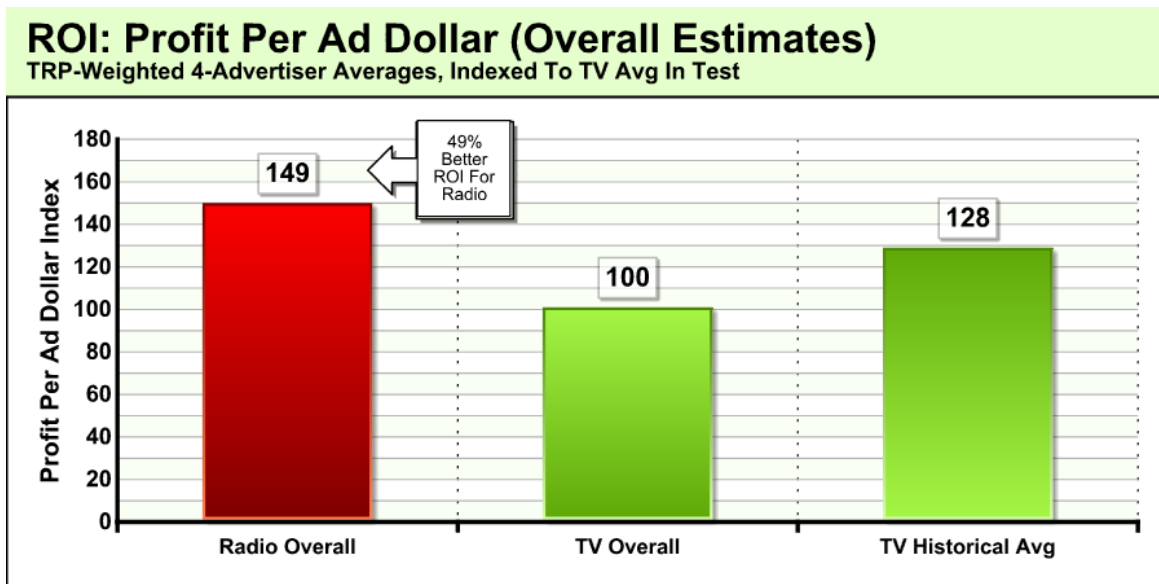
## Executive Summary

The Radio Ad Lab is pleased to present its third major new study in twelve months. This project is our largest yet, and it addresses the core issue of advertising—Return on Investment. The results confirm our prior theories: **Radio's ROI in this test was 49% higher than we observed for television.**

This “real world” study was conducted by Millward Brown and Information Resources, Inc. (IRI), and it examined four pairs of radio and television campaigns in a range of product categories over a six month period. The product categories included Grocery Food, Grocery Non-Food, and two very distinct Over-The-Counter Drug products.

By conducting the study in four IRI BehaviorScan markets, we were able to create four test cells—one with no TV or incremental radio; one with incremental radio only; one with national TV only; and, one with both national TV and incremental radio.

After the six month campaigns, we were able to analyze scanner sales data for each of those controlled test cells. The key finding is that **incremental radio campaigns showed significantly better ROI for these advertisers than did their national television campaigns**, whether the TV ROI was measured by this test's results or by the advertisers' own historical return estimates for television:<sup>1</sup>



<sup>1</sup> Our research partner, IRI, recommended that we benchmark our television test results to the advertisers' own internal estimates of TV ROI as reassurance that our TV test results were reasonable.

For all the study's complexity, we think the implications for advertising are clear:

- **Radio moves product.** Across four different advertisers, incremental radio advertising consistently and significantly increased product sales and delivered meaningful profit for each dollar of advertising. Radio demonstrated in this study that it can function as a primary medium for advertising.
- **Radio ads increase sales even when national television is present.** Radio was just as potent in the presence of 50–100 TRPs of national TV as it was by itself. In fact, the test results actually suggested slightly more impact for radio when combined with television than when used alone.
- **Radio's effects can be measured—when radio is used at sufficient weight.** Radio is prepared to be held accountable for its advertising effectiveness. But effectiveness measurement requires that advertising be present at sufficient weight for statistics to accurately capture that result.
- Most importantly: **Radio's ability to deliver strong Return on Investment for advertisers has been proven in a real-world test at last. In this study, radio's ROI was 49% better than television's,** and radio advertising's *value* is no longer just speculation.

One final thought: We can only guess how much better that value might be if the creative quality of radio advertising received as much attention and investment as ads in other media. All the television campaigns in this test had received favorable advance testing; none of the radio ads were pretested.

## Table of Contents

Executive Summary .....	2
Table of Contents .....	4
Introduction.....	5
Expectations.....	6
Origin Of The Current Project.....	7
The Test Design .....	7
The Participating Advertisers .....	9
Radio Ad Lab's Requirements .....	10
The Campaigns .....	11
The Sales Results .....	13
Sales Effects .....	13
Accounting for Investments (ROI) .....	17
ROI: The Simple Summary.....	20
Telephone Surveys: Directional Support.....	21
The Questions.....	21
Key Results .....	22
Discussion .....	23
Implications .....	24
Appreciation .....	26
Appendix A: Technical Appendix.....	27
IRI BehaviorScan Panel Methodology .....	27
Millward Brown Telephone Survey Protocol .....	31
Appendix B: Ad Schedule Miscellany .....	34
Television.....	34
Radio .....	35
Appendix C: The Imperial Margarine Study.....	36



## Introduction

Founded in 2001, the Radio Ad Lab is a nonprofit organization dedicated to creating and disseminating research about how radio advertising works in the U.S.

One of our distinctive attributes is the active participation of advertisers and agencies in the design of the research and in the communication of our results. That rare collaboration of buyers and sellers, combined with the serious financial commitment of the radio industry, has resulted in a series of significant new studies about radio ad effectiveness.

Our first primary research projects were published in 2004, and they provided significant new learning about how radio ads affect consumers, and about how radio can work with other media:

- In *Personal Relevance, Personal Connections: How Radio Ads Affect Consumers* (August 2004), conducted by WirthlinWorldwide, we found that the personal nature of the radio medium is also reflected in the perception that radio *advertising* is directed personally to the listener. That results in a powerful advertising environment of personal relevance and connections.
- In *The Benefits of Synergy: Moving Money Into Radio* (December 2004), conducted by The PreTesting Company, we learned how powerful radio can be when some ad dollars are shifted from television or newspapers into radio. Among other things, we found that swapping out one of two television ads for two radio ads increased unaided brand recall by 34%. Similarly, replacing one of two newspaper exposures with two radio ads almost tripled unaided brand recall.

Those studies were an excellent beginning, especially when combined with the knowledge from past research summarized in the Radio Ad Lab's series of *White Papers*, *Case Studies*, and the *Research Compendium*, all available for free at the Radio Ad Lab web site, <http://RadioAdLab.org>.

But the past research left us hungry for more. Specifically, we wanted to develop a deeper understanding of radio's actual Return on Investment (ROI) as an advertising medium, especially in a real-world setting.

## Expectations

Previously, there was only one recent North American study which explored radio's in-market sales effects publicly, captured in a 1999 paper by Sharon Paskowitz and Bill Bennett. Their paper, "A Test of Media Effectiveness for Imperial Margarine," was presented at the 1999 ESOMAR international radio conference ("Radio On The World Stage," Boston), and it described an in-market test of radio's direct effects on sales compared to television. They:



*...found no difference in sales in markets where equal weights of radio and TV advertising were used. However, a statistically significant higher volume of sales resulted when radio advertising ran at 50% higher weight levels than television.*

Appendix C contains a more detailed description of this study. Using in-market sales testing, it showed just how well radio ads can move product when radio is used at sufficient weight, and it showed that radio can be at least as effective as TV when the weights are comparable. This Canadian study served as a good model for the type of research we wanted to do in the United States.

In addition to the encouraging findings from that study, we had also seen a good deal of published research on *indirect* measures of effectiveness that showed how powerful radio can be when compared to television. As we state in our second *Radio Ad Lab White Paper*,<sup>2</sup> a review of about half a dozen studies of comparative recall and impact measures found that:

*Despite the variations in method, recency, and country, there is some convergent validity here. Assuming that recall is a reasonable proxy for effectiveness, it seems clear that Radio is about 80% as potent as TV—for a single exposure of an average spot.*

That, of course, *implies* a better ROI for radio, assuming that a radio spot costs less than 80% of a TV spot's price. But the Radio Ad Lab wanted to go further, toward direct measurement of ROI, for more products, and to do it in the U.S. We had good reason to *expect* a strong ROI for radio compared to television. But now we needed a robust study to examine that possibility in the real world.

---

<sup>2</sup> *Radio Ad Lab White Paper #2: Estimating the Benefit of Adding Radio* (Published May 2003), available at <http://radioadlab.org/whitepapers.html>.

## Origin Of The Current Project



To identify the best providers and designs for our research, the Radio Ad Lab Research Committee conducted a comprehensive search in Fall 2003 using a detailed Request for Proposals. That search yielded many good ideas which were winnowed to the three that received funding for our first round of new projects. The current study is the third of those original projects to be completed.

One of the themes that we stressed in the RFP was our strong interest in research on radio's ROI, along the lines of the Imperial Margarine study described above. Fortunately, the Millward Brown Company (MB), in partnership with Information Resources, Inc. (IRI), responded with the proposal that became the study we're reporting here.

Specifically, Millward Brown and IRI proposed that we use IRI's BehaviorScan test market capabilities to control the delivery of television commercials, while simultaneously using test and control *markets* to deliver or not deliver radio advertising. The concept called for a six-month test of actual campaigns for multiple advertisers.

The Radio Ad Lab Research Committee reviewed the initial concepts with MB and IRI in detail, and refined the concepts carefully. That led us to the following specific design.



## The Test Design

No real-world test can be perfectly controlled, especially over a six-month period, but we sought as much control as possible.

The control over television advertising exposure was provided by IRI's BehaviorScan Targetable TV capability in four of their test markets (Cedar Rapids, IA; Eau Claire, WI; Midland, TX; and Pittsfield, MA). In those markets, IRI has the ability to control whether specific TV ads are delivered to each household in each market's main cable system, and those cable systems account for a majority of the markets' TV viewers. (See the Technical Appendix for more details on IRI's BehaviorScan methodology.)

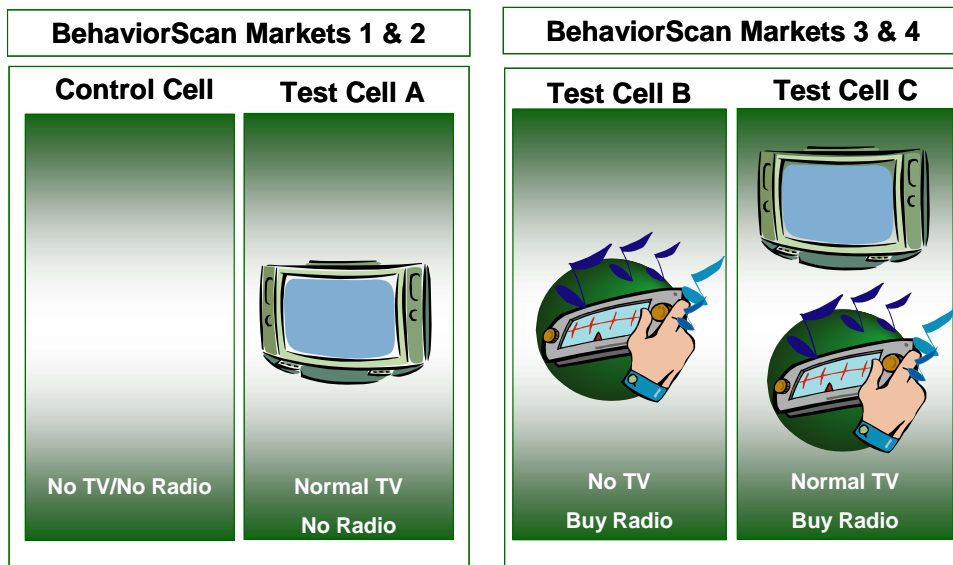
For our purposes, half of those households would receive the normal national TV advertising for the participating brands, and the other half would have those ads replaced with a public

service announcement. That substitution occurs real-time in IRI's studios for both broadcast and cable channels, and those cut-outs are performed successfully over 90% of the time.<sup>3</sup>



That means that half the cable households in these four markets received TV ad exposures for our participating advertisers, and half did not—a near-perfect test/control situation for television. But obviously, we didn't have the same capability for radio.

Therefore, to control exposure to radio advertising, we went with *market-level* controls: Two of the four markets had radio advertising purchased for this study and two did not. The specific markets with and without radio varied for each advertiser, depending on their sales profiles across markets; we tried to match the radio test and control markets appropriately for each advertiser. Here's how the original concept looked in graphical form:



As it happens—the “real world” being what it is—we weren't able to perfectly control radio exposures; several of our participating advertisers were regular users of national network radio, and that usage continued during our test. For those advertisers, then, our study became a test of *incremental* radio—of adding significant Target Rating Points (TRPs) in two of the four test markets, while the normal network radio remained in all four markets. However, our

<sup>3</sup> Actual cut-out effectiveness in our study ranged from 90% to 95%.

additional ad purchases outweighed the network radio on average by about four to one except for one of our advertisers, as discussed further below.

In all four markets, a subset of the cable households are also part of IRI's consumer shopping panel in which UPC-scannable purchases are recorded electronically at participating stores. This yields household-by-household information on product sales for correlation with media exposures. For the four participating advertisers with scannable products in grocery outlets, then, we were able to track differences in sales volume as the ad campaigns progressed, with the effects of television and incremental radio campaigns isolated.

Millward Brown brought their measurement expertise to bear with pre- and post-campaign telephone surveys among these cable households.<sup>4</sup> The surveys measured saliency, awareness, brand consideration, and brand behavior, among other things. We also asked each advertiser to propose one key question cluster to make sure that we measured the characteristics that most mattered to that advertiser and product type. (See the Technical Appendix for more details on the telephone protocols and sample outcomes.)

Finally, we depended on IRI's analytical experience to track and to control mathematically for other factors in the final analysis, including brand penetration, store-level variations, promotional activity, etc. This also included additional "panel matching" to control for market-to-market population variations. (Again, please see the Technical Appendix for more details on the statistical methods.)

Once we had the conceptual design in place, all we needed was a few brave advertisers...

## The Participating Advertisers

Doing "real world" research means using actual ad campaigns in real markets. That, of course, requires the cooperation of specific advertisers that are willing to modify their television campaigns in our test markets.

Since we wanted to work with several advertisers, it also involved significant coordination challenges. We'll spare you the details here, but you can imagine the logistical challenges needed to line up multiple advertisers for a single coordinated six-month test.



---

<sup>4</sup> Millward Brown conducted their surveys among the cable households which were *not* part of the IRI scanner panel, but which did have their TV commercial exposures controlled.

In the end, we're grateful to have received the cooperation and support of five advertisers in the following categories:

- Grocery Food
- Grocery Non-Food
- Over-the-Counter Drugs (two advertisers with very distinct products)
- Retail Chain

As we'll discuss in a moment, we ended up not being able to use the data for the retail chain in this study, but we're very appreciative of their cooperation.

The Radio Ad Lab granted the participating advertisers the exclusive rights to the detailed category-level data. And we agreed to maintain strict confidentiality about the specific advertisers and products tested.

In addition, we agreed that we will only present aggregated results in public (combining all four advertisers' data).

To gain the cooperation of advertisers, the Radio Ad Lab made a significant investment in this project, funding all the actual test costs. The advertisers then paid for the expense of radio advertising in two BehaviorScan markets; the cost of any new radio creative required; and in conjunction with the affected agency, any expenses associated with executing and monitoring the radio buys. Note that all the television advertising involved was already-planned national advertising; there was no local TV advertising present for any of our partners. See Appendix B for more details on the ad campaigns.

The testing of the ad campaigns began on or after September 27, 2004. By March 13, 2005, all test campaigns had ended.

## **Radio Ad Lab's Requirements**

Though we were anxious to gain advertiser cooperation, the Radio Ad Lab did have some requirements for participation in the study, including:

- **Duration:** The incremental radio campaigns had to air for as many weeks as the corresponding TV campaigns.

- **Weight:** The incremental radio campaigns had to be at least 100 TRPs per week if there was no network radio, or at least 50 extra TRPs if there was network radio in use.
- **Monitoring:** The advertisers or agencies had to commit to reasonable monitoring of radio station affidavits to assure that the test radio campaigns ran as scheduled.
- **Products:** The radio campaigns had to be for the same products with the same general themes as the TV campaigns.
- **Schedule Access:** The agencies needed to cooperate with IRI to provide scheduling information sufficient for managing the TV cut-outs.
- **Financials:** Campaign expense information was required in enough detail to allow IRI to compute relative ROIs. (This data was provided only to IRI, not to the Radio Ad Lab.)

You'll note that we did not require radio commercial pretesting for participants in the study, even though all of the television campaigns had received favorable testing in advance. More specifically, we did not require that the radio ads be demonstrated to be "good" ads by some pretesting measure.

We had two reasons for not making pretesting a requirement:

- As a practical matter, most of the advertisers expressing a willingness to participate in the study would not have been able to participate had this been a requirement.
- Furthermore, most of the advertisers being recruited were not routine users of radio ad testing services. Since the Radio Ad Lab has a policy of not recommending or endorsing methods of radio ad testing, it was not appropriate for us to direct these advertisers toward particular testing companies.<sup>5</sup>

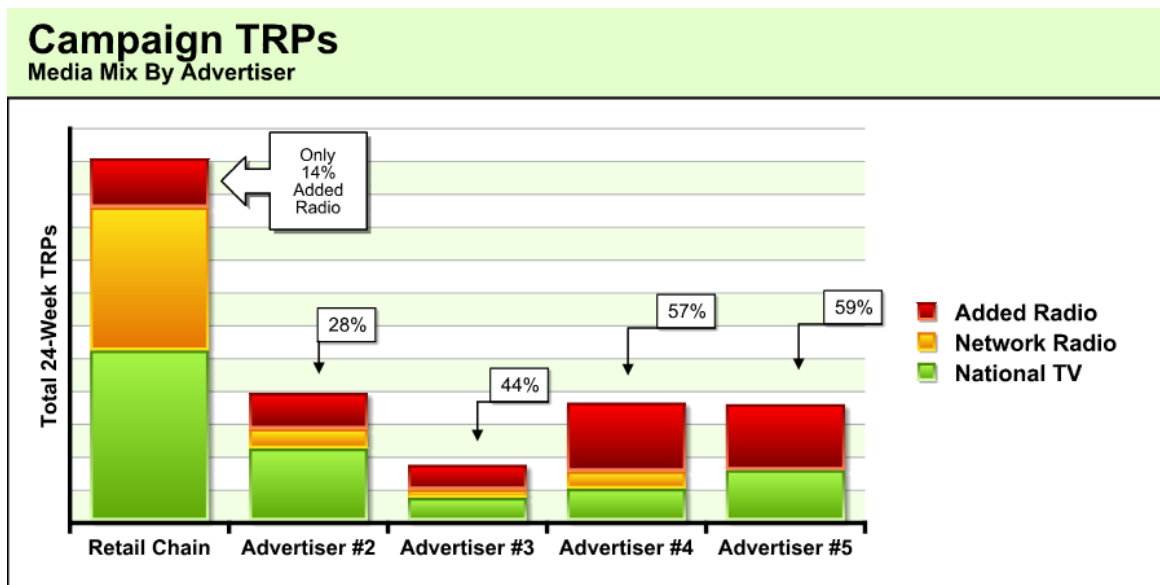
## The Campaigns

In setting guidelines for participation in this test, the Radio Ad Lab hoped to insure that our test would include additional radio advertising at sufficient weight to be clearly measurable, especially compared to the advertising weight being used for national television. Those guidelines worked well for all but one campaign.

---

<sup>5</sup> However, we do continue to recommend that interested parties review our *Guide to Commercial Testing Services for Radio*, available for free at <http://RadioAdLab.org>. This guide lists available information about a variety of testing providers without making specific recommendations about methodology.

In the end, the participating national retailer still had a test radio buy that was dwarfed by their other advertising, as can be seen in the chart below:



Clearly, our advance criteria didn't anticipate the overall level of national television and national radio advertising being used by this advertiser. Future studies like this will need to also consider *total* advertising weight, and how that total weight may affect the study's ability to parse media contributions.

In addition to this media weight issue, the retail campaign posed two other challenges for the Radio Ad Lab:

- This was the only advertiser for which we had no IRI scanner sales data. The only measures of effectiveness would have been the less sensitive pre-/post-campaign telephone survey measures. We knew this going into the study, of course, but not having direct sales measures further complicated the analytical challenges of the media mix issue.
- Finally, we did examine some of the telephone survey data, and perhaps not surprisingly, the results contained a number of anomalies. Among other things, the incremental radio and incremental radio + TV groups showed *lower* levels for awareness and usage than did the pre-wave groups—a counter-intuitive finding. The TV-alone findings were also surprisingly low, given the heavy amount of television weight used by this advertiser.

For all those reasons, we decided to exclude the retail chain from further analysis in this paper (though the retailer does have access to its own telephone data, as we promised).

Meanwhile, here are some other details on the campaigns for the other four advertisers:

	Advertiser #2	Advertiser #3	Advertiser #4	Advertiser #5
<b>Target</b>	M 18-49	A 25-49 (Q4) W 35-64 (Q1)	W 25-49	A 25-54
<b>Weeks</b>	22	17	21	24
<b>Existing Radio TRPs</b>	600	250	524	0
<b>Incremental Radio TRPs</b>	1100	750	2095	2000
<b>TV TRPs</b>	2198	705	983	1392

As you can see, we had a good mix of adult targets, and varying uses of radio and TV. More details on the ad schedules are available in Appendix B.

Note that all the radio ads were 30-second ads, as the participating advertisers were using existing network radio ads. Most of the TV ads were also 30 seconds; a few were 15 seconds.

Of course, we don't pretend that these are representative of all television or radio campaigns. In particular, we don't claim that these are typical of specific product categories. Our purpose here is simply to demonstrate that we had a reasonable array of different approaches to television and radio advertising so that our aggregation of results can be considered a good cross-section.

## The Sales Results

Though the study's design and analysis methods were complex, the results themselves are ultimately simple. We set out to answer a straightforward question:

- How does radio advertising's Return on Investment compare to television's?

We can now answer that question, in two steps. First, we estimate the in-market sales effects. Then we estimate how much it cost to achieve those effects.

### Sales Effects

The study's design permits us to look at the short-term (six-month) sales effects of radio and television in several different ways.



Remember that we had four distinct groups of consumers in the test, each receiving a unique combination of incremental radio and national television:



- 1) **No incremental radio, no television.** This group of people had no exposure to the incremental radio, and no exposure to the national TV ads. For three of the four advertisers, this group would have had exposure to the lower level of network radio ads, but that was true for all four test groups.



- 2) **Incremental radio, no television.** This group of people *did* have exposure to the incremental radio, but still no exposure to the national TV ads. They also had exposure to the same lower level of network radio ads as all the other groups.



- 3) **No incremental radio, with television.** This group did *not* have exposure to the incremental radio, but they *did* get exposed to the national TV ads for our advertisers. And they had the network radio ad exposures.



- 4) **Incremental radio, and television.** This group had exposure to *both* the incremental radio *and* the national TV ads, as well as the lower level of network radio ads.

The most interesting results come from comparing the six-month trends within these groups to each other. That's how we determine the effects of incremental radio in isolation, and of television in isolation.

Of course, we could just look at groups 2 and 3 over time to see those effects. But that doesn't consider what might have happened "naturally." For some brands, sales might be going up or down organically, seasonally, or as a result of other factors.

Furthermore, we're interested in seeing whether radio and TV effects differ when used alone compared to when they're used in combination with each other.

For all those reasons, we'll present our results (with a little simplifying shorthand) as follows:

**Radio In The Absence Of Television =**



Vs.



This results from the comparison of Group 2 to Group 1—of the group which *had* incremental radio (and no television) to the group which *did not have* incremental radio (and no television). This is (loosely) the effect of incremental radio by itself.

**Radio In The Presence Of Television =**



Vs.



This results from the comparison of Group 4 to Group 3—of the group which had incremental radio *and* television, to the group which *did not have* incremental radio but still had television. This is (loosely) the effect of incremental radio when added to national television—not the combined effect of radio and television, but rather, the incremental effect of *adding* radio to television.

**Television In The Absence Of Radio =**



Vs.



This represents the comparison of Group 3 to Group 1—of the group which had television with no incremental radio to the group which had no television and no incremental radio. This is the effect of national television by itself.

**Television In The Presence Of Radio =**



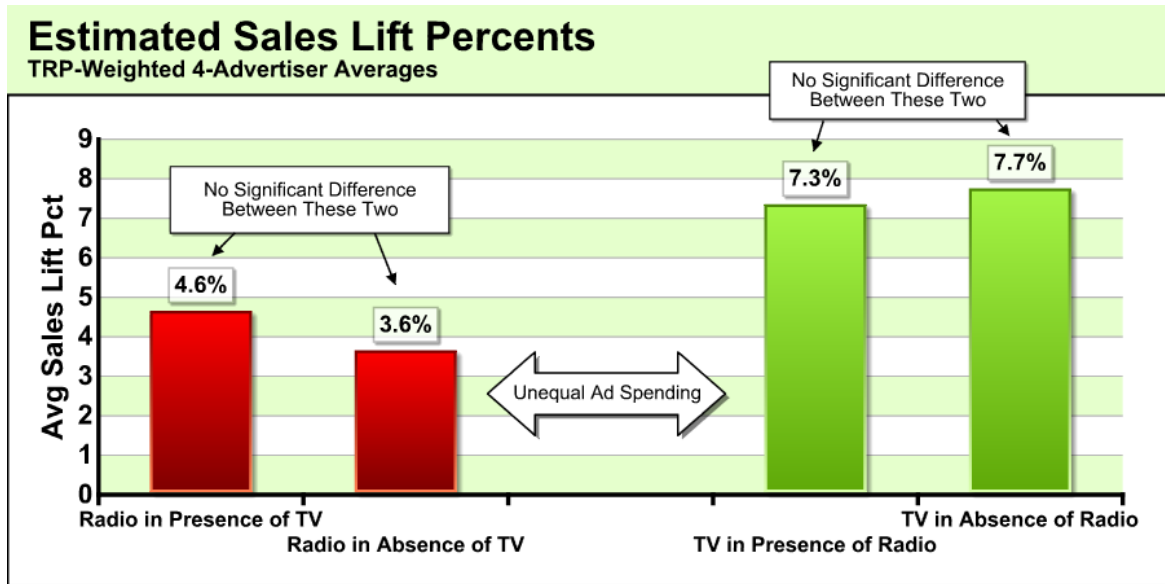
Vs.



This comes from the comparison of Group 4 to Group 2—of the group which had television and incremental radio to the group which had *no* television but which still had incremental radio. This is roughly the effect of *adding* national television to existing radio.

When IRI made those comparisons on the sales data, and controlled for all other external factors (see the Technical Appendix), here's what they saw. Remember that we haven't yet controlled for advertising weight or costs:<sup>6</sup>

<sup>6</sup> To compute the averages across four advertisers for this and subsequent charts, we weighted the individual advertiser results by the number of TRPs used by that advertiser in that medium across the duration of the study. Thus, heavier users of radio would contribute more to the radio average, and heavier users of TV would contribute more to the TV average.



As you can see, even before taking radio's lower costs into account, the radio ad campaigns for these four advertisers acquitted themselves very well. The incremental radio campaigns were linked to statistically significant sales lifts of 3.6% to 4.6% according to IRI, while the national TV campaigns were higher at 7.3% to 7.7%.<sup>7</sup>

There's some directional evidence of a synergy effect for radio, suggesting that radio works even better when it's added to television. The sales lift from incremental radio was slightly higher (4.6%) when the estimation was with groups that already had national TV campaigns present. The effect of radio in the *absence* of TV was slightly lower at 3.6%. However, the statistical difference between those two values is weak. In fact, it would be fair to average the two together to say that the sales lift from incremental radio in this study averaged 4.1%.

We do believe that radio can work especially well in combination with other media, as demonstrated in our own *Benefits of Synergy* study. But for this study and these campaigns, the synergy effects can't be isolated statistically, and we'll pool the effects from these groups for the purposes of simplifying our discussion of ROI.

Meanwhile, television showed a sort of reverse synergy, at least in direction. But again, the difference between TV in the presence of radio and TV in the absence of radio is not statistically significant. In the end, we would combine the two and say simply that the average sales lift from network television in this study was 7.5%.

<sup>7</sup> The three highest values above were statistically significant at the 90% confidence level or above. The lowest value of 3.6% was significant at the 86% confidence level.

Which leads to our first significant conclusion:

**At roughly similar weights when summed across all advertisers, incremental radio delivered more than half the short-term sales effect of national network television buys. Furthermore, radio's effect on sales was at least as large, if not larger, when television was already in use at a level of 50-100 TRPs.**

That's a somewhat lower radio-to-TV ratio than the other studies we cited at the beginning of this paper. But those studies were generally based on "per *exposure*" measures; in the present comparison using "weight," the TRP levels are based on national program estimates from TV peplemeters versus local station daypart estimates from radio diaries. Those aren't strictly comparable.<sup>8</sup>

We also need to remember that the television campaigns in this study were national buys for major advertisers. Given the pricing of national television, we might expect to see radio "catch up" when costs are taken into account.

### ***Accounting for Investments (ROI)***

Now that we have an estimate of short-term sales effectiveness—now that we see how incremental radio lifted product sales relative to national TV—we have to see how the relative *costs* of those ads compare.



In the real world, there are many ways to buy ad time on television and radio, of course, including national vs. spot and many permutations in between. For this study design, though, there was really only one way to make a clear comparison between radio and television costs.

In this test, the television ads were part of a national campaign, consisting of a variety of broadcast networks, cable networks, and syndication buys. The radio ads, conversely, were all bought locally in two of the four IRI BehaviorScan markets. Therefore, there were only two "actual" costs available to IRI in this project—the actual national TV costs paid by these advertisers, and the actual costs for the radio buys in the small markets of Cedar Rapids, Eau Claire, Midland, and Pittsfield. Somehow, that didn't seem like the right comparison to use for estimating ROI for the advertising industry.

---

<sup>8</sup> If nothing else, the two ratings systems have methodological differences that can affect the overall size of the rating points allocated to each spot.

Our solution was to “nationalize” the actual test radio buys into theoretical national radio campaigns. With the help of the Radio Ad Lab’s agency members that regularly purchase network radio advertising, we and IRI assumed that the incremental TRP levels used in this test were purchased for a national campaign on appropriate radio networks. Those estimated costs were then used by IRI for comparisons to the actual costs of the TV campaigns.

We had several reasons for doing so:

- We didn’t see a practical alternative. There’s simply no definitive way to come up with a national average for a *spot* radio campaign that would be proper to compare to the national TV campaigns in this study. Spot radio campaigns typically aren’t “national” anyway, and there are too many other variables (market rank, daypart, etc.) involved with spot pricing to compute an inarguable “average.”
- Conversely, and similarly, there was no acceptable way to translate the national TV campaign expenditures in this study into “typical” *spot* TV costs for comparing to the incremental radio buys in this study.
- All of the radio ads in this study were 30-second ads borrowed from existing network radio campaigns. That length is common in network radio, but not (yet) typical for local spot radio.
- Finally, the national television advertisers with which we partnered on this study would most likely be considering network radio for their radio investments. So this would be the right comparison for these advertisers.

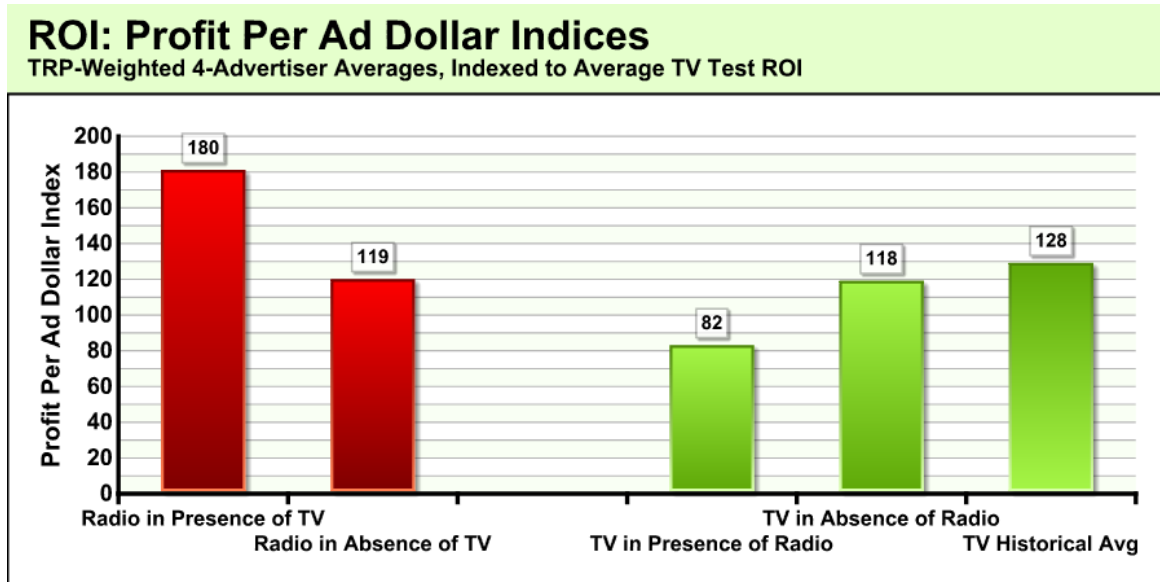
At the end of the day, the approach we took was the most appropriate for this study.

Keep in mind that we’re examining *total* Return On Investment in this analysis, not “incremental” ROI. Some ROI studies examine the return from each *additional* dollar of advertising, but our approach is to estimate the total return from *all* of the test advertising in each medium.

Also note that IRI computes ROI as estimated *profit* per dollar of advertising. This is derived by using IRI’s standard estimates of profitability by category—specifically, using an estimated average industry retailer margin of 30% and a manufacturer margin of 40%.

IRI strongly recommended that we present the ROI data as indices rather than absolute values. The underlying profit estimates are based on rough assumptions about margins, so the

precise dollar values are less important than the relationships across media.<sup>9</sup> Therefore, since our project is focused on radio's ROI *relative to television*, all the ROI indices presented here treat the four-advertiser average television ROI in the test as the baseline of 100.



Why is there a fifth bar on this chart? That's because IRI felt that the "TV in the presence of radio" estimate was not typical of their usual observations in past studies for these categories. To ensure maximum credibility for this study, they suggested that we seek historical estimates of television ROI from our participating partners for comparison purposes.

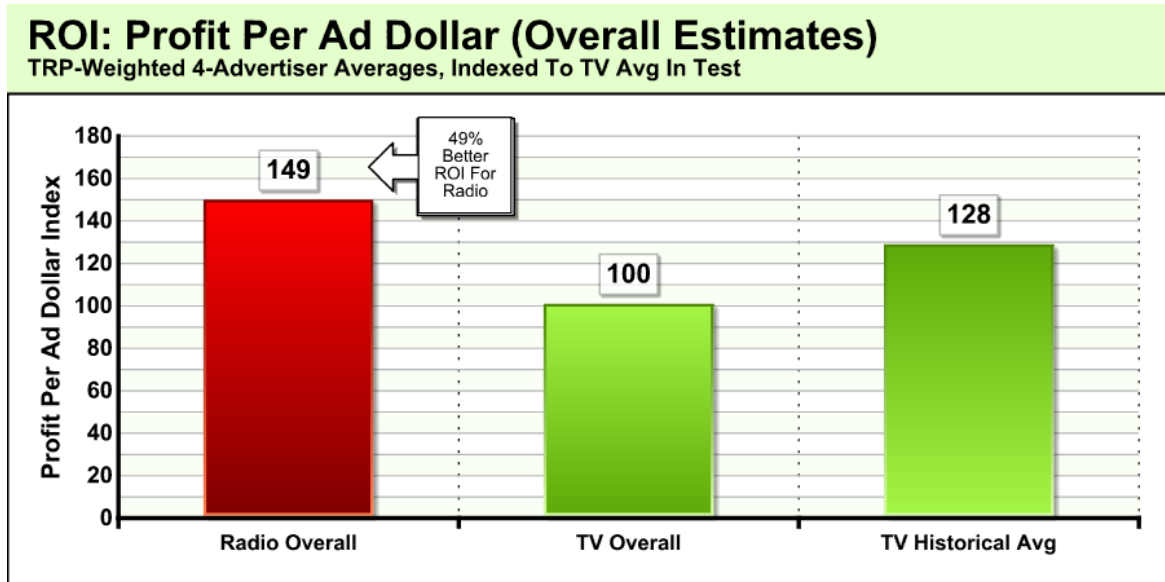
These historical estimates aren't strictly apples-to-apples comparisons as they represent annual (twelve-month) estimates of television return, while our study was a six-month test. Nevertheless, it provides another benchmark for consideration by readers.

Again, there's a hint of a synergy effect for radio here, in that the six-month ROI for radio in the presence of TV (i.e., the additional ROI from adding radio to the national TV campaign) is larger than the ROI for radio in the *absence* of an existing TV campaign. But because the sales-lift effects that we saw earlier are not considered statistically different, we should probably not make too much of this difference unless it reappears in future studies.

<sup>9</sup> However, we should point out that all of the underlying ROI values were in the range of \$0.28 to \$0.61 in estimated profit per dollar spent on advertising, so be assured that we're not using indices to inflate small differences.

Therefore, we believe it's appropriate to combine the radio results together and combine the television results together into a simple presentation. For the sake of completeness, we'll continue to present the TV historical average along with the TV test results, and we end up with the following very simple summary.

### ROI: The Simple Summary



And that's the bottom line. For these four advertisers, in this test, **incremental radio delivered 49% better Return on Investment than the corresponding national television campaigns**, using the best available comparative cost estimates.

## Telephone Surveys: Directional Support

In addition to the compelling in-market sales effect data presented above, we also sought guidance from pre- and post-campaign telephone surveys conducted by Millward Brown. Details on the survey methodology can be found in the Technical Appendix.

Because the results from the sales data were clear, and because we had to exclude the one advertiser whose results were to be measured exclusively with telephone survey data, we opted to present only a few of the telephone results here. In future analyses, we may revisit the telephone data set for additional learning, but for now, we look primarily for directional consistency with the IRI conclusions.

As described earlier, the telephone surveys were conducted among other households in the BehaviorScan cable television systems—those households whose TV ad exposures were controlled for this test, but whose shopping was not being tracked through IRI consumer panel membership. Separate samples of these households were surveyed before and after the campaign period.

We don't expect as much precision in these results; among other things, we have not controlled for extraneous marketing variables as we did in the sales analyses. These tabulations are based on simple random samples of adults 18-54 in the target populations, and no effort was made to match groups on purchase patterns.

That said, we were encouraged by the following results.

### ***The Questions***

In this report, we'll focus on three of Millward Brown's questions:

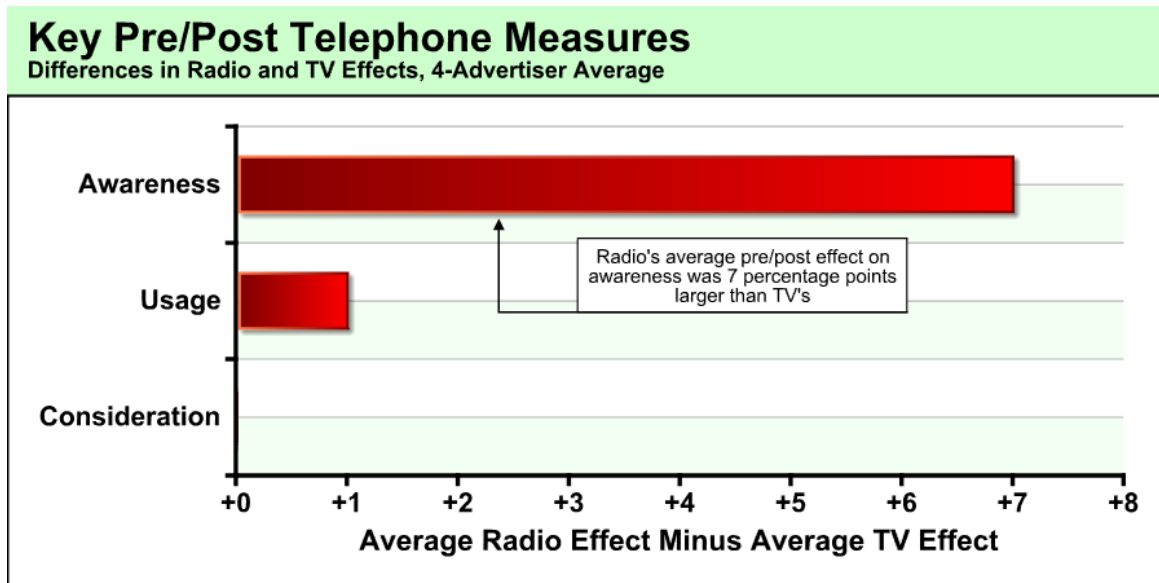
**Total Unaided Brand Awareness:** Typically, this question began with the statement, "Please think of brands of [product category]." That's followed by a one-sentence description of that product category, and then: "What brands of [category] have you seen or heard of? What other brands have you seen or heard of? What others?" Our tallies will count all mentions of the test brands.

**Use Most Often:** This question takes the form, "What one brand of [product category] do you use most often?"

**Consideration:** Here the respondent is asked, "How likely are you to consider [specific test brand] the next time you buy [product category]? Would it be..." The consumer then has several choices that include "your first choice," "one you would seriously consider," etc. We'll focus on the "first choice" responses.

## Key Results

Directionally, the pre-/post-campaign telephone survey results are also supportive of the positive results seen for radio in the sales data:



To keep the data presentation concise, we simply computed the average difference in pre-/post-campaign effects for incremental radio. We then contrasted that difference with the corresponding TV data for these three measures.

The statistically significant finding is that **incremental radio, on average, correlated with a seven-point larger impact on total unaided brand awareness than did the national television campaigns.** And that's without controlling for expenditure; these measures are simply unadjusted results for roughly similar radio and TV weight levels across the four advertisers.

We also saw roughly *equal* performance for incremental radio (compared to television) for the Usage and First-Choice Consideration measures.

We believe the strongest findings in this study are from the sales results presented in the previous section. But there's further support in the telephone measures for our general conclu-

sions. At significantly less cost, incremental radio delivered measurable marketing returns for these advertisers.

## Discussion

The Radio Ad Lab began this project with some reasonable expectations about radio's Return on Investment. While there aren't many studies that directly address the ROI equation, there are good reasons to expect that radio's relative effectiveness is greater than its relative cost when compared to television. We summarized several of those other studies at the beginning of this paper.

So the Radio Ad Lab isn't surprised that our new study demonstrated favorable ROI for radio. Though we took great pains (and made a considerable investment) to design an objective test, it seems quite logical that radio's sales effects would look powerful when considered in the context of total investment, especially relative to national television advertising.

With any single study, there are caveats, and this project is no exception. These are four products measured for six months at one point in time. These advertisers are all well-respected national marketers, but they aren't necessarily typical even of their own categories. A different mix of advertisers could well yield different results.

Furthermore, three of the four advertisers were already regular users of radio, as evidenced by their use of network radio during the study period. Perhaps these advertisers were more experienced with the medium (though that played no role in their recruitment).

That said, however, none of the advertisers did any pre-testing of their radio creative. Three of the four campaigns used radio ads that were direct derivatives of the TV creative, and that may or may not have been the most effective way to use the radio medium for these brands. (See, for example, our third *White Paper* and our first major primary research project, *Personal Relevance, Personal Connections*, for more about the unique communication characteristics of effective radio ads.)

These brands were all well established brands with a history of using national television advertising regularly. Decay (or "adstock") from previous television advertising certainly spilled into the test period; conversely, longer-term effects of the radio advertising were *not* captured. A longer-term test could turn out differently. That also means that a test with new brands might yield different results.

Of course, these products were all of the “CPG” category—scannable consumer packaged goods. That was necessary to achieve the powerful household-level sales tracking required by this study design, but it's at least a theoretical limit on generalizability.

As noted earlier, we had to make assumptions about ad pricing to compare television to radio. Readers may choose to make informal adjustments to apply these results to different applications (e.g., spot TV or spot radio).

So, yes, there are limits on the projectability of a study like this. But we also believe this is a major milestone in learning about radio advertising.

The Radio Ad Lab is considering now whether to continue this concept into additional tests so that we can keep building this knowledge base. But for now, we believe this first-ever large-scale public test of radio's ROI in the U.S. has much to teach us.

---

## Implications

For all the study's complexity, we think the implications for advertising are clear:

- **Radio moves product.** Across four different advertisers, incremental radio advertising consistently and significantly increased product sales and delivered meaningful profit for each dollar of advertising. In this study, radio demonstrated that it can function as a primary medium for advertising.
- **Radio ads increase sales even when national television is present.** Radio was just as potent in the presence of 50–100 TRPs of national TV weight as it was by itself. In fact, the test results actually suggested slightly more impact for radio when combined with TV than when used alone.
- **Radio's effects can be measured—when radio is used at sufficient weight.** Radio is prepared to be held accountable for its advertising effectiveness. But effectiveness measurement requires that advertising be present at sufficient weight for statistics to accurately capture that result.
- Most importantly: **Radio's ability to deliver strong Return on Investment for advertisers has been proven in a real-world test at last. In this study, radio's ROI was 49% better than television's,** and radio advertising's *value* is no longer just speculation.

One final thought: We can only guess how much better that value might be if the creative quality of radio advertising received as much attention and investment as ads in other media. All the television campaigns in this test had received favorable advance testing; none of the radio ads were pretested.

## Appreciation

First, and most importantly, the Radio Ad Lab is grateful to the advertisers that worked with us on this project. Our only regret is that we can't acknowledge them more specifically. We have to hope that they received as much value from the exercise as we did, and we are very, very appreciative of their cooperation.

In addition, we thank Millward Brown and IRI which brought this study concept to us in the first place, and which executed the research so professionally. More specifically, we'd like to acknowledge the key project team members who gave both expertise and many long hours to this study:

Millward Brown

Philip Herr

David Hluska

Jennifer Davis

Information Resources, Inc.

J. P. Beauchamp

Amy Hess

Neal Heffernan

Of course, we're grateful to the members of the Radio Ad Lab Research Committee for their (volunteered) time and expertise for this and all of our projects. They approved and refined the project originally; they helped in many logistical details along the way; and of course, they reviewed the full study results and approved this summary. The current membership of this nonpareil group of advertiser, agency, and broadcast researchers is listed at the Radio Ad Lab website at <http://RadioAdLab.org/committee.html>.

The funders and Board of the Radio Ad Lab should receive special acknowledgement for this study. This project was a huge leap of faith for them, representing a major investment with no certainty that this very public project would yield a favorable return on *their* investment. The current Radio Ad Lab Board is listed at <http://RadioAdLab.org/board.html>.

Finally: the Radio Ad Lab would like to acknowledge the contributions of research consultant Jim Peacock of Peacock Research, Inc. His direction, insight and guidance have been a critical asset to this project. Among other things, Mr. Peacock is the author of this report.

## Appendix A: Technical Appendix

### ***IRI BehaviorScan Panel Methodology***

IRI's BehaviorScan is a system of five geographically dispersed test markets carefully selected to provide geographic representation and reasonably typical consumer demographics and retail conditions. In each market, a panel of approximately 3,000 households makes it possible to track purchasing behavior on an item-by-item level over time.

In four of these markets—the four used by the Radio Ad Lab—the BehaviorScan within-market split-cable Targetable TV system delivers different TV advertising to different households within the same market.

Part of the preparatory work for all BehaviorScan tests involves completing a feasibility analysis to insure proper test design and representivity for the product. IRI determines market selection for a test based on both consumer and store information to include category and brand development indices, panel-to-store coverage and demographic representation to insure test readability. (This is how markets were paired for the presence and absence of radio advertising in the Radio Ad Lab test.)

### **Household Panel/Data Collection**

The BehaviorScan panel is comprised of approximately 3,000 households within each of five markets:

- Eau Claire, Wisconsin
- Cedar Rapids, Iowa
- Pittsfield, Massachusetts
- Grand Junction, Colorado (not used in the Radio Ad Lab test)
- and Midland, Texas.

These markets are intended to be representative of the Total U.S., small enough to control, and affordable for both media buys and new product production quantities. On-going panel recruitment techniques seek to maintain panels that are representative of the population.

IRI employs a multi-step recruitment process which includes:

- a geographically dispersed (within the desired geography) direct mail invitation to join the Shoppers' Hotline panel
- a follow-up telephone call to describe details of the program and gain panelist commitment
- equipment sent within 24 hours to accepted panelists
- a follow-up welcome call one week after shipment

Each active panelist receives a monthly newsletter, *Key Issues*. In addition to providing scanning tips, answering panelists' questions, providing informative articles and printing panelists' recipes, the newsletter announces the winners of monthly drawings and quarterly sweepstakes.

The monthly drawings are generally for cash prizes, while the quarterly sweepstakes offer theme-related prizes such as computer systems, exercise equipment, etc. Panelists are also guaranteed prizes for their active participation through the Shoppers' Choice point-based incentive program. Participation allows panelists to earn points that can be redeemed for gift certificates at local retailers, movie theaters, restaurants or through mail-order catalogs.

BehaviorScan data collection is done at the household level. Retail sales and consumer purchase data can be captured in nearly all supermarkets, drug stores, and mass merchandisers in each market to allow for faster and more comprehensive evaluations of new products. (To accommodate the software complexities of the Radio Ad Lab's test design, only grocery store data were used for the Radio Ad Lab test.)

	SUPERMARKET	DRUG	MASS	OTHER
CEDAR RAPIDS, IOWA	S, C, P	S, C, P	S, P	CONTACT IRI FOR OPTIONS
EAU CLAIRE, WISCONSIN	S, C, P	S, C, P	S, P	
MIDLAND, TEXAS	S, C	S, C	S	
PITTSFIELD, MASSACHUSETTS	S, C	S, C	S	
GRAND JUNCTION, COLORADO*	S, C	S, C	S	

S= POINT OF SCALE SCANNER DATA; MASS EXCLUDES WAL-MART  
 C=HOUSEHOLD PANELISTS USE CARDS AT POINT OF SALES AT PARTICIPATING STORES  
 P= PANELISTS SCAN AT HOME, CAPTURING PURCHASES MADE AT NON-PARTICIPATING STORES  
 \*TARGETABLE TELEVISION IS NOT AVAILABLE

A sub-set of panelists in two BehaviorScan markets, Cedar Rapids and Eau Claire, also use a handheld scan device at home to capture purchases made outside of participating BehaviorScan stores, but because that data is only available in two markets, that information was not used for the Radio Ad Lab test.

### **Analysis of Test Effect**

IRI's Targetable TV system delivers different advertising to up to three pre-assigned groups of demographically-balanced households within each selected market; the Radio Ad Lab test used a two-group design. For each particular test, the groups are then further matched to ensure an "apples to apples" comparison. A statistical linear procedure forms groups of IRI panelist households with similar quarterly category and brand purchasing and demographics.

Utilizing IRI's panel match program, two equal panel groups were created for the Radio Ad Lab test within each BehaviorScan test market based on data from the 52 week Pre-Period. The following match variables were used to pick the "Test" and "Control" cells:

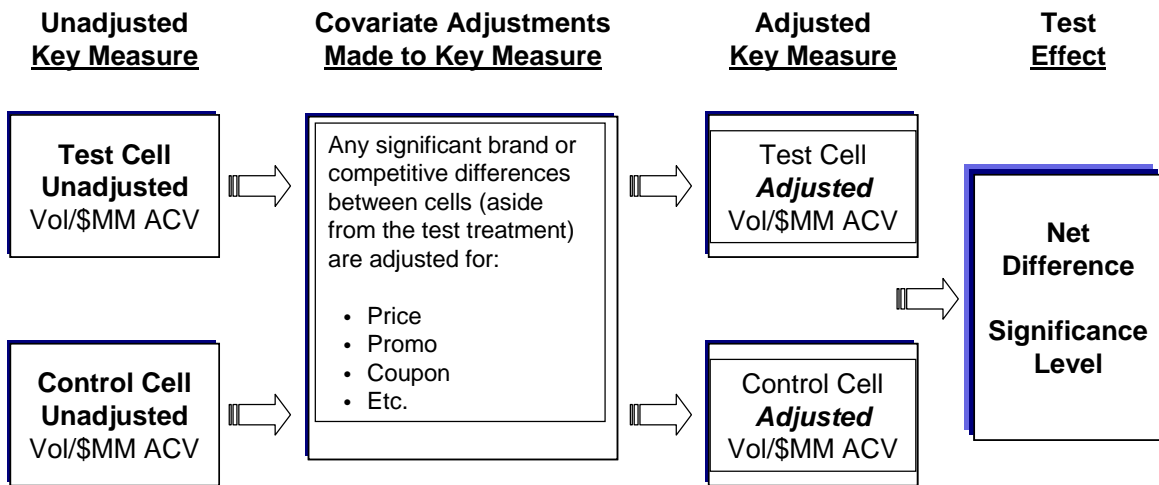
- Test Brand Volume/1000 HHs
- Test Brand Penetration (% HHs Buying)
- Sample Category Volume/1000 HHs
- Sample Category Penetration (% HHs Buying)
- Competitor A Volume/1000 HHs
- Competitor A Penetration (% HHs Buying)
- All Other Competitors Volume/1000 HHs
- All Other Competitors Penetration (% HHs Buying)

At the completion of test execution, IRI's statistical analysis of test results determines the net change in sales attributable to the test variable, after adjusting for any extraneous factors that may also have impacted sales. In the Radio Ad Lab test, the analysis specifically controlled for variances across groups in the following other factors:

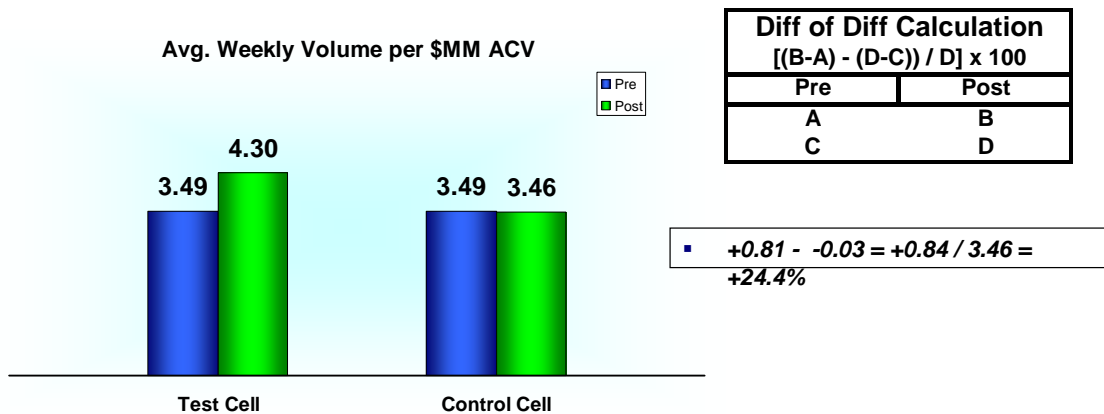
- "Test Brand" Price
- "Test Brand" Feature Only
- "Test Brand" Display Only
- "Test Brand" Feature & Display
- "Test Brand" Price Reduction Only
- "Test Brand" Average Items Carried
- Competitive Price
- Competitive Trade Promotion

Competitive Average Items Carried  
Category Trend

This analysis of covariance (ANCOVA) is a statistical procedure that identifies, measures, and adjusts for the influence of activities not directly related to the advertising that occurred during the test period. This allows IRI to isolate the true test effect from other factors.



The test impact is analyzed through the Net Effect calculation. The Net Effect compares the difference between the pre-test and test periods in the Test Cell to the same difference in the Control Cell. By calculating a Net Effect, we take into account both the changes that would have occurred in the absence of the test, as well as those during the test, to get a total impact due to the test scenario. For example:



For more technical details, please contact the Radio Ad Lab at [info@RadioAdLab.org](mailto:info@RadioAdLab.org).

## **Millward Brown Telephone Survey Protocol**

This section concerns the telephone surveys conducted by Millward Brown before and after the ad campaigns to assess effects on awareness and other indirect measures.

### **Sample Plan**

IRI (via Survey Sampling, Inc.) provided the telephone sample to Millward Brown. The sample consisted of cable household telephone numbers from within the BehaviorScan markets excluding IRI scanner-panel participants.<sup>10</sup>

Each piece of sample was tagged with which cell the sample was in for each test. For example, household number 1 may be in the TV-present, incremental radio cell for brand number 1 but no TV, no radio for brand number 2. This was specified on the sample record.

IRI provided the entire list of potential respondents to Millward Brown before the benchmark waves. Because of the timing of the ad campaigns, MB actually conducted interviews in three waves—two waves at the beginning, and one combined wave at the end. The first pre-wave was conducted for three advertisers, and shortly thereafter, the second pre-wave was conducted for the final two advertisers. The post-campaign wave included all five products. Sample was allocated equally and randomly across all three waves.

Because we had a relatively small amount of sample available in the Pittsfield market, we opted to add a pre-survey warm-up letter with a \$2 incentive in that market to optimize response rates there. That sample shortage also required that we cycle some incomplete sample from the early waves into the post-campaign wave.

### **Questionnaire Structure**

A single random respondent within each household was selected for interviewing via the “next birthday method.” No substitutions were allowed.

The questionnaire was structured so that every respondent received the same set of questions. Each respondent was asked screener qualification questions including age (18-54), commercial radio listenership (two hours or more per week) and no critical industry affiliation. Criti-

---

<sup>10</sup> Within the BehaviorScan markets, IRI maintains a consumer shopper panel. This panel is a subset of the households with controlled ad delivery, and the scanner panel is made up of about 2,000 households within each market (each market has between 16,000 and 45,000 households). The households in the IRI consumer shopper panel were not included in the sample that IRI provided to Millward Brown, but those households *were* included in the analysis conducted by IRI for sales tracking purposes.

cal industries included ad agencies, marketing research companies and any companies that make or distribute the brands of the categories included in the test.

Once the respondent passed through the screener, they were asked a series of questions regarding one of the five categories included in the test. Questions include unaided brand awareness, aided brand awareness, whether they were the primary decision maker for the category, usage (ever tried, past six month and most often used brand), future purchase consideration, total brand communication awareness, and for the post wave only, aided impressions of the radio advertising and ad recognition/brand identification of specific radio executions. Advertiser-supplied questions were also included here.

Once respondents completed these questions for one category, they began another loop of these questions for a different category. The order of the categories was randomized.

Brand awareness questions were asked of key brands in the category. The remainder of the questions was asked about the specific test brand. Within the usage questions there were one or two category usage questions.

Once a respondent answered questions for all five categories involved in the test, they were asked briefly about media consumption habits. Finally, respondents were asked classification questions such as marital status, household size, employment, income, gender and ethnicity.

The interviews for the shorter pre-campaign calls averaged about 19 minutes in length. The post-campaign calls (which included all five products) averaged about 29 minutes.

## **Data Collection**

Millward Brown call centers dial seven days a week. Interviewing occurred from 4PM to 9PM on weekdays (respondent time). Saturday hours were 9AM to 5PM and Sunday hours were 1PM to 8PM. Based on the Radio Ad Lab requirements, Millward Brown used dialing algorithms which yielded twelve or more dialing attempts on non-contacts. In addition, the project-specific algorithms allowed for daypart dialing ensuring that all unresolved sample received at least one weekday attempt, six evening attempts and three weekend attempts.

Refusal conversion attempts were made after about a week of "cooling off" time. The best MB interviewers were assigned to conversion attempts, with different types of refusals (initial refusals, hang up refusals, selected respondent refusals, etc.) being sent to different interviewers. A five dollar supplemental premium was offered as an incentive to cooperate.

Bilingual interviewers from Millward Brown's bilingual call center were available for this project and the questionnaire was translated into Spanish. Any telephone number in any market where the respondent indicated that they did not speak English or would prefer to do the interview in Spanish was sent to the bilingual call center. The Midland market made up the bulk of these calls as it is roughly 40% Hispanic.

Every interviewer was monitored and given feedback. Supervisors monitored a minimum of 15% of each interviewer's completes daily. Additionally, a minimum of 10% of screeners were monitored daily to verify the respondent was properly screened and qualified to participate in the study.

All interviewers received approximately 15-18 hours of interviewer training.

### Sample Disposition Data

We estimate the response rate over all three waves of telephone interviewing to have been 19.3%, computed as follows:

7,234	Completed interviews
2,255	Qualified during screener, but refused interview
[7,246]	[Not qualified but completed screener]
	[Estimated 57% qualification rate among screener cooperators]
21,316	57% of refusals that refused prior to the screening question
6,723	57% of usable noncontacts (minus obvious unusables like disconnects)
19.3%	$= 7,234 \div (7,234 + 2,255 + 21,316 + 6,723)$

## Appendix B: Ad Schedule Miscellany

In addition to the information in the main body of the report, we can provide the following additional information about the ad campaigns that were scheduled for this study.

### **Television**

The television ad campaigns for the Radio Ad Lab's partner advertisers (which were the normal and already-planned national TV campaigns for these advertisers) utilized a wide array of dayparts and networks:

#### **Advertiser Targeting Women 25-49**

- Prime 10%
- Day 30%
- Syndication 10%
- Early Morning 10%
- Cable 40%

#### **Advertiser Targeting Men 18-49**

- Prime 15%
- Late Night 6%
- Day Cable 6%
- Cable News 6%
- Prime Cable 23%
- Sports 44%

#### **Advertiser Targeting Adults 25-54**

- Prime 20%
- Day 20%
- Early Morning 5%
- Late Night 10%
- Syndication 5%
- Cable 40%

#### **Advertiser Targeting Adults 25-49/Women 35-64**

- Prime 30%
- Day 33%
- Cable 34%
- Early News 3%

## **Radio**

The radio spot schedules were generally daytime run-of-schedule, utilizing a mix of AM Drive, PM Drive, Mid-Day and weekend. The advertiser targeting Men 18-49 allocated about 40% of their spending towards sports programming, which is consistent with what was done on television.

## Appendix C: The Imperial Margarine Study

Part of our inspiration for the current study was provided by Sharon Paskowitz and Bill Bennett, authors of a paper presented at the 1999 ESOMAR international radio conference ("Radio On The World Stage," Boston). Their paper, "A Test of Media Effectiveness for Imperial Margarine," described an in-market test of radio's actual effects on sales compared to television, and they...



"...found no difference in sales in markets where equal weights of radio and TV advertising were used. However, a statistically significant higher volume of sales resulted when radio advertising ran at 50% higher weight levels than television."

For this study, Imperial Margarine funded the purchase of the TV and radio advertising and the cost of the creative; BBM Canada funded the research itself; Ogilvy and Mather was the ad agency, while Harrison, Young, Pesonen and Newell purchased media time. Grocery store scanners collected product sales data, and the Modeling and Analytics Group of AC-Nielsen in Canada conducted the analysis of those sales.

The Winnipeg market acted as the control market, receiving 1200 TRPs<sup>11</sup> of TV time across nine weeks. The radio test markets were Calgary and Edmonton, with Calgary receiving 1200 TRPs of radio time during the same nine weeks, and Edmonton receiving 1800 TRPs in radio. As was typical for Imperial, there was no national TV.

In addition to tracking Imperial margarine sales via scanner data, the project also utilized pre- and post-study awareness telephone surveys conducted by ComQUEST, a division of BBM Canada. The telephone part of the study focused on brand and commercial recall.

Key results from the analysis included these observations:

- Actual sales were about the same when equal amounts of advertising weight in both radio and television markets were applied.
- A 50% higher weight of radio advertising provided almost 11% more sales than the TV-only market.

---

<sup>11</sup> That was an increase from the typical television levels usually used by Imperial, for reasons explained in the Paskowitz/Bennett paper.

- But the pre- and post-advertising brand awareness measurement was “quite equivocal” in its results, perhaps because Imperial’s awareness was already very high.

As a side note, this study also confirmed an observation seen in numerous other studies—that consumers who had no opportunity to see an ad on television (i.e., the only ads they could have heard were on radio) still tended to attribute their ad exposures to TV.

Obviously, this study showed powerful results about radio’s relative strength for moving product, but the authors were reluctant to make a direct translation into statements about Return on Investment:

“No cost-per-rating point information was revealed during this test, so we cannot comment on the relative strength of radio versus television when comparing the media on a cost-to-actual-sales basis. However, in Canada a radio GRP is about half the cost of a television GRP for the same target group. In fact, Ogilvy & Mather has told us that to buy TV at the same weight level as radio in the test markets would have cost double the cost of radio. Therefore, it would be an easy leap in logic to say that the radio medium can deliver the same or more sales as TV but at a lower cost. However, there are many factors that should be taken into account before this universal statement can be made.”

Overall, then, this Canadian study contributed significantly to the body of knowledge about radio ad effectiveness. Using actual in-market sales testing, it showed just how well radio ads can actually move product when radio is used at sufficient weight, and it demonstrated that radio can be at least as effective as TV when the weights are comparable.