Marketing Accountability Standards

Practices & Processes Underlying the Development & Management of an “Ideal Metric”

Posting for Industry Feedback

October 2010

Marketing Accountability Standards Board of the Marketing Accountability Foundation
Introduction

*Measuring and Improving the Return from TV Advertising (An Example)* was selected, conducted & reviewed by members of The Boardroom Project (Oct 06 – July 07), reviewed by MASB Directors (Aug 07 – March 08) and posted to the MASB website for industry feedback (April 08).

Feedback concerning the TV Project included “Would be more useful if the measure and provider were identified”, and “The metrics piece is confusing – call it ARS persuasion”.

In Feb 08, comScore announced the acquisition of the ARS Group and in March 08, MASB Directors determined it was time to name the TV “Exemplar” measure for clarity and context...at least for the current project which documents the Practices and Processes Underlying the Development and Management of an “Ideal Metric”.

TV Example “Ideal Metric” Identified

Based on a specific *behavioral* measure of *consumer brand preference*

\[
persuasion_{\text{ARS Persuasion}} \quad \text{APM Facts*} \quad \frac{\text{Change In Consumer}}{\text{Brand Preference}} \quad \text{(Choice)} \quad \text{Market Results}
\]

* APM Facts = ARS Persuasion Metric for ads that actually air versus the same methodology used at other stages of the advertising development process.

Source: “Measuring and Improving the Return from TV Advertising (An Example),” MASB, April 2008
Summary & Conclusions (TV Project)

The (ARS Persuasion/APM Facts) consumer Brand Preference Metric has met the MASB Marketing Metric Audit Protocol (MMAP).

Its characteristics would deem it “ideal” for serving as a standard for measuring and forecasting the impact of TV advertising and for managing and improving the return.

Application of the metric during the advertising development and management processes has enabled improvement in return greater than that needed to offset the rises in TV Media costs.

Note: While various metrics may be called the same and even look alike in many ways, specific methodologies within classes and types of metrics often yield very different levels of reliability and validity (see Appendix B)

Source: “Measuring and Improving the Return from TV Advertising (An Example),” MASB, April 2008
The Practices & Processes
Underlying the Development & Management
of ARS Persuasion/APM Facts…an “Ideal Metric”
Measurement is The Foundation for Reporting, Forecasting & Improving Return

**Measurement**
- Reliably identify business opportunities (or threats) given current context & (potential) actions (MR Vision 2003); Process of achieving & maintaining measurement reliability, predictive validity, sensitivity & calibration.

**Near Sight**
- Profound understanding (of the business process or human & customer behavior) that yields a clear prioritization of action; Learning or principles that yield true predictions with unvarying uniformity (IBID); Process of explaining variance/identifying the causal drivers of the business or human behavior.

**Insight**
- Documented method of operating that yields higher level of performance than other operating behaviors (IBID); Process of applying Knowledge to the operating process for improved performance.

**Foresight**
- Analytical technique that represent causal relationships among various conditions & actions taken to achieve specific business results, and forecast future outcomes of various potential actions & conditions (IBID)

Source: “The Improvement Pyramid,” MASB Project Agenda, Work in Process
Corporate Practices (ARS Group)

- Triumvirate in operating structure
  - Research
  - Operations
  - Marketing

- Research structure
  - Measurement Standards
    - Reliability, validity, etc
    - Integrated data bases
  - Knowledge
    - Identification of drivers (causals)
    - Better Practice implications & application

- Customer Support
  - Teaming w/Marketing (inc customer service)
  - Customer specific research-on-research
  - 30% of activity was self-funded research (on-research)
MMAP: 10 Characteristics of an Ideal Metric

1. Relevant . . . addresses specific pending action
2. Predictive . . . accurately predicts outcome of pending action
3. Objective . . . not subject to personal interpretation
4. Calibrated . . . means the same across conditions & cultures
5. Reliable . . . dependable & stable over time
6. Sensitive . . . identifies meaningful differences in outcomes
7. Simple . . . uncomplicated meaning & implications clear
8. Causal . . . course of action leads to improvement
9. Transparent . . . subject to independent audit
10. Quality Assured . . . formal/on-going process to assure 1–9
Practices: Measurement Standards & Knowledge (ARS Group)

- Behavioral nature of methodology
  - Relevant (#1)
  - Objective (#3)
  - Simple (#7)
- Systematic test-retest reliability process (ARSAR) (#10)
  - Reliable (#5)
  - Sensitive (#6)
- Systematic validity process (ARSAV) (#10)
  - Predictive (#2)
  - Calibrated (#4)
  - Sensitive (#6)
- On-going better practice insight process (BPI)
  - Causal (#8)
- Documentation, publication, academic audits & collaboration
  - Transparent (#9)
Behavioral, Relevant, Objective, Simple
ARS Brand Preference (choice) Methodology

QUICK SERVE HAMBURGER RESTAURANTS
($5.00 CERTIFICATE)
Please circle the ONE you truly want

- A&W
- Burger King
- DQ
- Hardee's
- Sonic
- Wendy's

YOUR FAVORITE OTHER HAMBURGER RESTAURANT (write in)

Behavioral, Relevant, Objective, Simple
Practices: Measurement Standards & Knowledge

- Behavioral nature of methodology
  - Relevant (#1)
  - Objective (#3)
  - Simple (#7)
- Systematic test-retest reliability process (ARSAR)
  - Reliable (#5)
  - Sensitive (#6)
- Systematic validity process (ARSAV)
  - Predictive (#2)
  - Calibrated (#4)
  - Sensitive (#6)
- On-going better practice insight process (BPI)
  - Causal (#8)
- Documentation, publication, academic audits & collaboration
  - Transparent (#9)
Reliability Process (ARSAR)

- Process managed by “Standards” research team
- Systematic test-retesting
- All key measures
- Continuous monitoring
- Feed to management (dashboard)
- Collaboration w/Operations when required
  - Operating process improvement
- Summary & publication every other year or so
Excerpted with Permission

Global Reliability of the ARS Persuasion® Measure, ARS® Validated Drivers, and Diagnostic Measures

February 2005 & February 2008 Documents

The ARS® Group

rsc THE QUALITY MEASUREMENT COMPANY
110 WALNUT STREET EVANSVILLE, INDIANA 47708 NA
TEL: (812) 425-4562 FAX: (812) 425-2844

©2005 rsc. All rights reserved.
The following list of rsc trademarks is provided for your information only. This is not a complete list of every service mark owned by rsc. Each mark, whether indicated to be a trademark in this document or not, is protected by state, national, or international law.
rsc trademarks are On The Global Playing Field and Now That's A Global Idea. rsc service marks include, but are not limited to, the following: APM, ARS, the ARS logo, ARS Context, ARS StarBoard, ARS WOWWW, ARS WOWWW CI, ARS Persuasion, ARSAR, C2C, CATS, Fair Share, Firstep, Outlook, PPD, Persuasion Points Delivered, StarBoard, Television Preview, VOC, and Window On the World Wide Web. The ARS method, process, and application/software for forecasting advertising's impact on sales (as determined by Market Mix Modeling output) are the subject of one or more pending patents.
Publication is ~45 pages and includes

- Background & Objectives
- Design/Data Collection (Database)
- Empirical Findings
- Business Implications & Applications
- Study Data and Inclusion Criteria (Appendix)
Context

Classical measurement theory holds that observed variability is the result of sampling error and “other” error as defined by the F-ratio\(^1\):

\[
\text{Observed Error} = \text{Sampling Error} + \text{“Other” Error}
\]

Sampling variability imposes known limits on the reliability of all sampling-based measures. The presence of “other” error variance would increase the observed measurement variability beyond that expected from sampling alone and decrease reliability.

\(^1\) See Appendix of Technical Remarks Regarding Statistical Analyses Used
### Analysis/Findings

**ARS Persuasion Measure**  
Test for Inconsistent ("Other") Error Variance

<table>
<thead>
<tr>
<th>ARS Persuasion Measure&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Commercial Test-Retest Pairs</td>
</tr>
<tr>
<td>Test-Retest Variation Observed&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>Variation Expected From Random Samples&lt;sup&gt;2&lt;/sup&gt;</td>
</tr>
<tr>
<td>F-Ratio</td>
</tr>
<tr>
<td>Conclusion</td>
</tr>
</tbody>
</table>

Globally, the ARS Persuasion measure is as reliable as the laws of random sampling allow. "Other" error variance is not a factor.

---

<sup>1</sup> Includes testing in Europe, Latin America, and North America for the time period January 2005 to December 2007.

<sup>2</sup> In standard deviation units.

<sup>3</sup> At the 95 percent confidence level.
## Analysis/Findings

### ARS Persuasion Measure

Chi-Square “Goodness-of-Fit” Analysis

<table>
<thead>
<tr>
<th>t-Value</th>
<th>Expected</th>
<th>Actual Percent of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percent of Cases</td>
<td></td>
</tr>
<tr>
<td>0.000–0.674</td>
<td>50%</td>
<td>52%</td>
</tr>
<tr>
<td>0.675–0.842</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>0.843–1.036</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1.037–1.282</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>1.283–1.644</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>1.645–1.959</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>&gt; 1.960</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

\[
\chi^2 \text{ at 90 Percent Cutoff} = 11.07 \quad \chi^2 \text{ at 95 Percent Cutoff} = 12.59
\]

<table>
<thead>
<tr>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.30</td>
</tr>
<tr>
<td>2.35</td>
</tr>
</tbody>
</table>

---

For the combined data set, the distribution of test-retest t-values does not differ significantly from what is expected.

---

1. See Appendix A
2. Includes testing in Europe, Latin America, and North America for the time period October 2005 to September 2007.
3. N = 256
Empirical Findings

ARS Validated Drivers of the ARS Persuasion Measure¹

<table>
<thead>
<tr>
<th>Positive Content Elements</th>
<th>Criterion</th>
<th>Test-Retest Agreement²,³</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brand Differentiation</td>
<td>Presence</td>
<td>99%</td>
</tr>
<tr>
<td>New Product/New Feature</td>
<td>Presence</td>
<td>100</td>
</tr>
<tr>
<td>Product Convenience⁴</td>
<td>Presence</td>
<td>99</td>
</tr>
<tr>
<td>Competitive Comparison</td>
<td>Presence</td>
<td>99</td>
</tr>
<tr>
<td>Superiority Claim⁴</td>
<td>Presence</td>
<td>99</td>
</tr>
<tr>
<td>Brand Name Reinforces Benefit⁵</td>
<td>Presence</td>
<td>100</td>
</tr>
<tr>
<td>Demonstration of Product in Use⁴</td>
<td>Presence</td>
<td>100</td>
</tr>
<tr>
<td>Setting Directly Related to Use⁴,⁵</td>
<td>Presence</td>
<td>100</td>
</tr>
<tr>
<td>Time Actual Product on Screen⁴,⁵</td>
<td>Time</td>
<td>100</td>
</tr>
<tr>
<td>Number of Brand Name Mentions⁴</td>
<td>Count</td>
<td>100</td>
</tr>
<tr>
<td>Time Brand Name/Logo on Screen</td>
<td>Time</td>
<td>98</td>
</tr>
<tr>
<td>Time Until Category Identified</td>
<td>Time</td>
<td>99</td>
</tr>
<tr>
<td>Time Until Product/Package Shown</td>
<td>Time</td>
<td>100</td>
</tr>
</tbody>
</table>

These positive elements have been validated in The ARS Group’s databases as helping to explain superior and below average ARS Persuasion outcomes relative to the Fair Share degree-of-difficulty benchmark and on absolute levels. All exhibit sufficient test-retest correspondence for use as drivers of the ARS Persuasion measure.

¹ See: “Summary of Factors That Differentiate Television Commercials That Are More or Less Sales Effective and the Business Implications and Applications of This Knowledge”; The ARS Group, December 2003.
² Percent of time the test and retest measures give the same “indication.”
³ N=144
⁴ Element also ARS Validated Driver of ARS Related Recall.
⁵ Element also ARS Validated Driver of ARS Key Message Communication.
Business Implications (Better Practices)

Globally, the sales-related *ARS Persuasion* measure is **behaving as reliably as random sampling will allow** and is free of extraneous sources of variance. It is therefore **appropriate to use standard t-tests to assess the significance of differences** when analyzing *ARS Persuasion* results on an operational basis.

The secondary measures validated as predictive of *ARS Persuasion* outcomes and the consumer feedback elements have high enough test-retest agreement to be useful in advertising development and in the explanation of test results.
<table>
<thead>
<tr>
<th>Region</th>
<th>Case Number</th>
<th>ARS Persuasion Level (Original Test)</th>
<th>Standard Error</th>
<th>ARS Persuasion Level (Re-Test)</th>
<th>Standard Error</th>
<th>t-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>US 1</td>
<td>7.6</td>
<td>1.78</td>
<td>0.8</td>
<td>1.72</td>
<td>2.74*</td>
<td></td>
</tr>
<tr>
<td>US 2</td>
<td>6.1</td>
<td>1.02</td>
<td>2.3</td>
<td>0.87</td>
<td>2.79*</td>
<td></td>
</tr>
<tr>
<td>US 3</td>
<td>3.9</td>
<td>0.81</td>
<td>3.1</td>
<td>0.77</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>US 4</td>
<td>0.9</td>
<td>1.50</td>
<td>2.2</td>
<td>1.21</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>US 5</td>
<td>18.6</td>
<td>1.86</td>
<td>16.8</td>
<td>1.72</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>US 6</td>
<td>2.4</td>
<td>1.70</td>
<td>-1.3</td>
<td>1.98</td>
<td>1.42</td>
<td></td>
</tr>
<tr>
<td>US 7</td>
<td>3.1</td>
<td>1.07</td>
<td>3.0</td>
<td>1.49</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>US 8</td>
<td>1.9</td>
<td>0.73</td>
<td>1.8</td>
<td>0.80</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>US 9</td>
<td>8.6</td>
<td>1.40</td>
<td>11.3</td>
<td>1.42</td>
<td>1.33</td>
<td></td>
</tr>
<tr>
<td>US 10</td>
<td>2.6</td>
<td>1.73</td>
<td>1.2</td>
<td>1.51</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>US 11</td>
<td>10.7</td>
<td>2.01</td>
<td>9.9</td>
<td>1.68</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>US 12</td>
<td>0.2</td>
<td>0.94</td>
<td>-1.2</td>
<td>1.06</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td>US 13</td>
<td>25.8</td>
<td>2.00</td>
<td>20.4</td>
<td>2.04</td>
<td>1.87*</td>
<td></td>
</tr>
<tr>
<td>US 14</td>
<td>3.2</td>
<td>0.80</td>
<td>2.8</td>
<td>0.85</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>US 15</td>
<td>11.2</td>
<td>2.10</td>
<td>11.8</td>
<td>1.90</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>US 16</td>
<td>7.4</td>
<td>1.61</td>
<td>7.2</td>
<td>1.62</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>US 17</td>
<td>1.7</td>
<td>1.56</td>
<td>2.8</td>
<td>1.66</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>US 18</td>
<td>6.4</td>
<td>1.32</td>
<td>5.6</td>
<td>1.36</td>
<td>0.42</td>
<td></td>
</tr>
<tr>
<td>US 19</td>
<td>2.1</td>
<td>0.85</td>
<td>1.6</td>
<td>1.12</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>US 20</td>
<td>2.8</td>
<td>1.80</td>
<td>2.6</td>
<td>1.47</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>US 21</td>
<td>3.4</td>
<td>1.04</td>
<td>1.4</td>
<td>0.89</td>
<td>1.48</td>
<td></td>
</tr>
<tr>
<td>US 22</td>
<td>1.8</td>
<td>0.58</td>
<td>3.2</td>
<td>0.74</td>
<td>1.48</td>
<td></td>
</tr>
<tr>
<td>US 23</td>
<td>15.0</td>
<td>2.13</td>
<td>12.9</td>
<td>2.41</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>US 24</td>
<td>3.2</td>
<td>1.31</td>
<td>7.2</td>
<td>1.46</td>
<td>2.96*</td>
<td></td>
</tr>
<tr>
<td>US 25</td>
<td>5.0</td>
<td>1.41</td>
<td>8.0</td>
<td>1.89</td>
<td>1.31</td>
<td></td>
</tr>
<tr>
<td>US 26</td>
<td>3.4</td>
<td>1.23</td>
<td>4.3</td>
<td>1.35</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>US 27</td>
<td>2.8</td>
<td>1.13</td>
<td>1.8</td>
<td>1.09</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>US 28</td>
<td>11.3</td>
<td>1.81</td>
<td>11.2</td>
<td>1.67</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>US 29</td>
<td>4.3</td>
<td>1.86</td>
<td>5.2</td>
<td>1.93</td>
<td>0.34</td>
<td></td>
</tr>
<tr>
<td>US 30</td>
<td>2.5</td>
<td>1.40</td>
<td>0.7</td>
<td>1.06</td>
<td>1.03</td>
<td></td>
</tr>
<tr>
<td>US 31</td>
<td>3.7</td>
<td>1.57</td>
<td>1.1</td>
<td>1.29</td>
<td>1.29</td>
<td></td>
</tr>
<tr>
<td>US 32</td>
<td>0.4</td>
<td>0.16</td>
<td>0.8</td>
<td>0.28</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>US 33</td>
<td>-0.4</td>
<td>0.31</td>
<td>0.4</td>
<td>0.47</td>
<td>1.38</td>
<td></td>
</tr>
<tr>
<td>US 34</td>
<td>3.8</td>
<td>0.93</td>
<td>5.1</td>
<td>1.01</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>US 35</td>
<td>5.1</td>
<td>1.02</td>
<td>4.1</td>
<td>0.92</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>US 36</td>
<td>7.0</td>
<td>1.48</td>
<td>7.6</td>
<td>1.45</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>US 37</td>
<td>3.3</td>
<td>1.72</td>
<td>2.0</td>
<td>1.61</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>US 38</td>
<td>4.8</td>
<td>1.08</td>
<td>3.1</td>
<td>1.13</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>US 39</td>
<td>1.4</td>
<td>1.15</td>
<td>0.5</td>
<td>1.02</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>US 40</td>
<td>4.2</td>
<td>1.53</td>
<td>2.7</td>
<td>1.40</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>US 41</td>
<td>0.0</td>
<td>2.88</td>
<td>-2.6</td>
<td>2.89</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>US 42</td>
<td>1.2</td>
<td>1.48</td>
<td>1.2</td>
<td>1.23</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>US 43</td>
<td>1.0</td>
<td>0.94</td>
<td>1.2</td>
<td>1.09</td>
<td>0.14</td>
<td></td>
</tr>
</tbody>
</table>

* Significant difference at 90 percent confidence level.
** Individual region codes: US = United States; CD = Canada, LA = Latin America; EU = Europe.
Practices: Measurement Standards & Knowledge

- Behavioral nature of methodology
  - Relevant (#1)
  - Objective (#3)
  - Simple (#7)
- Systematic test-retest reliability process (ARSAR)
  - Reliable (#5)
  - Sensitive (#6)
- Systematic validity process (ARSAV)
  - Predictive (#2)
  - Calibrated (#4)
  - Sensitive (#6)
- On-going better practice insight process (BPI)
  - Causal (#8)
- Documentation, publication, academic audits & collaboration
  - Transparent (#9)
Validity Process (ARSAV)

- Process also managed by “Standards” research team
- Tracking of all customer ads airing (eg Monitor Plus)
- Matching to ads tested
- Obtaining market data (eg third party Nielsen, IRI, IMS)
- Other data from customers (eg GRPs & Marketing Mix Results)
- Continuous monitoring
- Feed to management (dashboard)
- Collaboration with Marketing (Customer Service) when required
  - Test design process improvement (eg Lysol)
- Summary & publication every other year or so
- Journal publication of key learning (eg wearout)
Excerpted with Permission

Summary of The ARS Group’s Global Validation and Business Implications

November 2005 & January 2008 Documents

The ARS® Group

rsc THE QUALITY MEASUREMENT COMPANY

110 Walnut Street, Evansville, Indiana 47708
TEL: (812) 425-4562 FAX: (812) 425-2844

©2005 rsc. All rights reserved.

The following list of rsc trademarks is provided for your information only. This is not a complete list of every service mark owned by rsc. Each mark, whether indicated to be a trademark in this document or not, is protected by state, national, or international law.

rsc service marks include, but are not limited to, the following: APM, ARS, the ARS logo, ARS Context, ARS StarBoard, ARS WOWWW, ARS WOWWW CI, ARS Persuasion, ARSAR, Brand Preference Monitor, BP, BP Monitor, C2C, CATS, Fair Share, Firstep, HeartZone, HZ, MindZone, MZ, Outlook, PPD, Persuasion Points Delivered, rsc Brand Preference, Preference Points Delivered, StarBoard, Television Preview, VOC, and Window On the World Wide Web. The rsc method, process, and application/software for forecasting advertising’s impact on sales (as determined by Market Mix Modeling output) and for tracking consumer brand preference over time are the subject of two or more pending patents.

ADSAM is a registered mark of ADSAM Marketing, LLC.
Publication is ~116 pages and includes

Outline

• Background & Objectives
• Design/Data Collection (Database)
• Analysis
• Empirical Findings/Knowledge (Better Practice Insight)
• Business Implications (Better Practices)
• Business Applications (Best Practice Tools)
• Frequently Asked Questions/Additional Findings
• Study Data and Inclusion Rules (Appendix)
The relationship between *APM* Facts and in-market effects remains strong and is calibrated across brands, categories, and countries. A “2” is a “2” and a “7” is a “7” in the USA, Mexico, Germany, etc.

In fact, the strength of television advertising alone, as measured via the *ARS* methodology, explains over 50 percent of the variance in market share change across brands, categories, and countries. To our knowledge, no other single variable in the marketing mix has been shown to have such a dramatic influence over a 4-week period of time.

---

1 Source: R. L. Polk New Vehicle Registration, IMS HEALTH, IRI InfoScan, Markettrack, Nielsen SCANTRACK, or Nielsen Retail Index.
The strong relationship between APM Facts and market share change (r = 0.72) continues when the analysis period is expanded over a quarter.

Improving APM Facts just “two points” is associated with a +0.4 market share improvement over a quarter.

1 Source: R. L. Polk New Vehicle Registration, IMS HEALTH, IRI InfoScan, Markettrack, Nielsen SCANTRACK, or Nielsen Retail Index.
2 Quarterly share data were only available for 320 cases in the overall data set.
Summary/Conclusions
Variance Explained in Market Share Change (Over a Quarter)

<table>
<thead>
<tr>
<th></th>
<th>r</th>
<th>r^2</th>
<th>F Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRPs alone</td>
<td>0.26</td>
<td>0.07</td>
<td>95%</td>
</tr>
<tr>
<td><em>APM</em> Facts alone</td>
<td>0.71</td>
<td>0.51</td>
<td>99%+</td>
</tr>
<tr>
<td><em>APM</em> Facts &amp; GRPs</td>
<td>0.77</td>
<td>0.59</td>
<td>99%+</td>
</tr>
<tr>
<td><em>APM</em> Facts, GRPs, &amp; flighting</td>
<td>0.80</td>
<td>0.64</td>
<td>99%+</td>
</tr>
<tr>
<td><em>APM</em> Facts, GRPs, flighting, and Normal Competitive Environment</td>
<td>0.81</td>
<td>0.66</td>
<td>99%+</td>
</tr>
<tr>
<td><em>APM</em> Facts, GRPs, flighting, Normal Competitive Environment, Price, and Distribution</td>
<td>0.83</td>
<td>0.69</td>
<td>99%+</td>
</tr>
<tr>
<td><em>APM</em> Facts, GRPs, flighting, Normal Competitive Environment, Price, Distribution, and Error in Sales Data</td>
<td>0.84</td>
<td>0.71</td>
<td>99%+</td>
</tr>
</tbody>
</table>
Taking into account the branding activities explored so far, a total of ~70 percent of the variance in market share change is accounted for over a 12-week (quarter) period of time.

The ~65 percent representing TV activity demonstrates the power of TV advertising in the marketing arena.

It also recognizes the contribution of the advertiser and agency team in developing brands/strategies/ads that sell, i.e., “a couple of points” improvement in APM Facts yields +0.4 points in market share over a quarter.

While the three percent variance explained by Product Price and Distribution is small across all cases, the impact on individual new products with a price greater than 20 percent of the category average and/or distribution less than 40 percent can be substantial.

1Among total “free market” data set.
2Distribution includes new product cases only. Price is factored in only when the advertised product price is 20%+ greater than the category average.
Frequently Asked Questions/Additional Findings

I. Do the findings hold for:

A. Established and new brands?
B. Large and small brands?
C. Different regions?
D. Different types of products?
E. 30" and 15" ads?
F. Rational and emotional ads?
G. Advertised products and the total product line?
H. Different relationships to the *Fair Share* benchmark?
I. Samples containing men and women and women only?
J. Category purchasers?
K. Multiple purchase categories?
L. Seasonal categories?
M. Ads containing a competitive comparison?
N. Frames showing multiple product packages?
O. Categories with a strong store brand presence?
P. Old Pre/New Post frame methodology?
Q. “Halo” Effects?
R. Bi-monthly data?
S. A longer time period?
T. The most recent data set?
I. B. 1. Do the Findings Hold for Large and Small\textsuperscript{1} Established Brands?

Large and small brand cases show a relationship similar to that of the overall data set.\textsuperscript{3}

\textsuperscript{1} Large brands are defined as brands with a pre choice of 20% or greater; small brands have a pre choice of less than 20%.
\textsuperscript{2} Source: R. L. Polk New Vehicle Registration, IMS HEALTH, IRI InfoScan, Markettrack, Nielsen SCANTRACK, or Nielsen Retail Index.
\textsuperscript{3} At the 95% confidence level based on “Tests of Difference of Data Sets” as described under Design/Data Collection.
I. B. 2. Do the Findings Hold for Very Large\textsuperscript{1} Established Brands?

A strong relationship exists between advertising effectiveness and market share change among very large brands.\textsuperscript{3}

\textsuperscript{1} Very large brands are defined as brands with a pre choice of 40% or greater.
\textsuperscript{2} Source: R. L. Polk New Vehicle Registration, IMS HEALTH, IRI InfoScan, Markettrack, Nielsen SCANTRACK, or Nielsen Retail Index.
\textsuperscript{3} At the 95\% confidence level based on “Tests of Difference of Data Sets” as described under Design/Data Collection.
Frequently Asked Questions/Additional Findings

II. Additional Findings

A. Is airing an APM Fact of zero the same as going dark?

B. Does airing low(er) APM Facts (less than 2.0) produce greater sales effects than not advertising/going dark?

C. Does continuing to air low(er) APM Facts (less than 2.0) produce greater sales effects than going dark?

D. What percentage of tested ads score zero or less?

E. Does the share decline when going dark vary by category?

F. Can the share point decline when going dark be verified with an independent data set?

G. Why is using the PPD metric a more accurate and actionable method than using just GRPs or impact per 100 GRPs?

H. Is ARS Related Recall predictive of market results?

I. What protocols are applied for understanding ad impact within the context of other marketing activities?

J. How do APM Facts relate to ad impact when isolated from other elements of the marketing mix?

K. Has TV APM Facts and GRPs (PPDs) continued to explain similar variance in market share change over time?

L. Is the ARS Persuasion measure reliable?
II. A. 3. Is Airing an *APM* Fact of Zero the Same as Going Dark?

For Zero Level *APM* Facts

|                        | Going to Air | Continue to Air | Going Dark | Best Fit Market Share Change for the Typical Category
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Best Fit Market</td>
<td>- 0.4</td>
<td>- 0.4</td>
<td>- 0.4</td>
<td></td>
</tr>
<tr>
<td>Share Change(^1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the market share change model, going to air in the typical category with an ad having an *APM* Fact of 0.0 and continuing to air an ad that has worn down to an *APM* Fact of 0.0 have the same sales impact as going dark over the course of a quarter. This converges with the findings from our lab.

\(^1\)Based on the typical, moderately elastic category. Categories with greater elasticity will experience greater drops and categories with lower elasticity will experience lower drops.
Frequently Asked Questions/Additional Findings

II. Additional Findings

A. Is airing an APM Fact of zero the same as going dark?

B. Does airing low(er) APM Facts (less than 2.0) produce greater sales effects than not advertising-going dark?

C. Does continuing to air low(er) APM Facts (less than 2.0) produce greater sales effects than going dark?

D. What percentage of tested ads score zero or less?

E. Does the share decline when going dark vary by category?

F. Can the share point decline when going dark be verified with an independent data set?

G. Why is using the PPD metric a more accurate and actionable method than using just GRPs or impact per 100 GRPs?

H. Is ARS Related Recall predictive of market results?

I. What protocols are applied for understanding ad impact within the context of other marketing activities?

J. How do APM Facts relate to ad impact when isolated from other elements of the marketing mix?

K. Has TV APM Facts and GRPs (PPDs) continued to explain similar variance in market share change over time?

L. Is the ARS Persuasion measure reliable?
How do APM Facts relate to ad impact when isolated from other elements of the marketing mix?

**PPD Metric Predicts the Sales Impact**
*(Free Market—Isolated Impact From Market Mix Modeling)*

The relationship between the Persuasion Points Delivered metric and Market Mix Modeling outcomes data has a strong 0.90 correlation among the 112 cases in this data set.

1 Days of incremental category volume. Calculated by dividing Incremental Volume from TV advertising by average category volume per day.

2 With Normal Competitive Environment
The power of the ad accounts for most of the overall variation in TV impact as derived independently by Marketing Mix Modelers; metrics of GRPs for media weight, (APM Facts) for the TV messages, the wearout function and normal competitive environment function explain ~90% of variance.

* Marketing Mix Modeling Output: Sales Volume Impacted from TV.

Source: “Measuring and Improving the Return from TV Advertising (An Example),” MASB April 2008
Summary of Validity

Correlation with Trial/Volume/Share

1970s New Product *Reported* Trial (isolated impact) \[ r = +.85 \]
1980s Split-cable Copy Tests (isolated impact) \[ (7/7) \]
1990s Split-cable Weight Tests (isolated impact) \[ r = +.90 \]
2000s Marketing Mix Modeling Output (isolated impact) \[ r = +.91 \]
2000s Scanner Share Change (non-isolated impact) \[ r = +.72 \]

. . . (ARS Persuasion) predicts TV advertising’s impact on market results at ~.90 level when the TV activity is isolated from other elements of the marketing mix (about as high a relationship as possible, given sampling probability);

And at the ~.70 level within the context of other marketing activities (demonstrating the relative leverage of TV in the marketing mix, as well as the precision of this consumer brand preference/choice methodology).

Source: “Measuring and Improving the Return from TV Advertising (An Example),” MASB, April 2008
What are the Financial Implications of Precision?

(ARS Persuasion) detects about 2 points as significant at the 90% level of confidence . . . and a 2 point difference in results (airing just one ad) is associated with a .04 difference in market share over a business quarter.

In a category with sales of say $500M per quarter using just one ad scoring 2 points higher returns ~$2M more in sales for the same media costs and multiple ads return even more.

Source: “Measuring and Improving the Return from TV Advertising (An Example),” MASB, April 2008
Practices: Measurement Standards & Knowledge

- Behavioral nature of methodology
  - Relevant (#1)
  - Objective (#3)
  - Simple (#7)
- Systematic test-retest reliability process (ARSAR)
  - Reliable (#5)
  - Sensitive (#6)
- Systematic validity process (ARSAV)
  - Predictive (#2)
  - Calibrated (#4)
  - Sensitive (#6)
- On-going better practice insight process (BPI)
  - Causal (#8)
- Documentation, publication, academic audits & collaboration
  - Transparent (#9)
Better Practice Insight Process (BPI)

- Process managed by “Insight” research team
- Continual search for “drivers”
  - Drivers are things that can be acted upon
  - As opposed to traditional use of “diagnostics”
- Collaboration with Operations
  - Experiments to find additional drivers (hypothesis testing)
- Collaboration with Marketing (inc Customer Service)
  - Listening for hypothesis generation
- Summary & publication every other year or so
- Journal publication of key learning (eg brand differentiation)
Excerpted with Permission

Summary of Factors Affecting ARS Persuasion Scores

March 2005 & July 2007 Documents

The ARS® Group

©2005 rsc. All rights reserved.
rsc service marks include, but are not limited to, the following: APM, ARS, the ARS logo, ARS Context, ARS StarBoard, ARS WOWWW, ARS WOWWW CI, ARS Persuasion, ARSAR, C2C, CATS, Fair Share, Firstep, Outlook, PPD, Persuasion Points Delivered, StarBoard, Television Preview, VOC, and Window On the World Wide Web. The ARS method, process, and application/software for forecasting advertising's impact on sales (as determined by Market Mix Modeling output) are the subject of one or more pending patents.
Outline

Publication is ~156 pages and includes

- Background & Objectives
- Design/Data Collection
- Analysis/Finding
- Business Implications (Better Practice Insight and ROI)
- Business Applications (Best Practice Tools)
- Frequently Asked Questions
- Additional Findings/Other Questions
- Future Investigations
- Study Data and Inclusion Criteria (Appendix)
The several hundred conditions and elements explored over 3 decades explain 85 percent of the total variation in (ARS Persuasion) outcomes.

Source: “Measuring and Improving the Return from TV Advertising (An Example)” MASB April 2008
Practices: Measurement Standards & Knowledge

- Behavioral nature of methodology
  - Relevant (#1)
  - Objective (#3)
  - Simple (#7)
- Systematic test-retest reliability process (ARSAR)
  - Reliable (#5)
  - Sensitive (#6)
- Systematic validity process (ARSAV)
  - Predictive (#2)
  - Calibrated (#4)
  - Sensitive (#6)
- On-going better practice insight process (BPI)
  - Causal (#8)
- Documentation, publication, academic audits & collaboration
  - Transparent (#9)
Transparency

- Stewart & Furse audit (1980’s)
- JP Jones audit (1990’s)
- MASB audit (2000’s)
- Hundreds of publications by multiple firm members
- Dozens of journal papers
  - With customers (eg Adams, Schroiff)
  - With academics (eg Stewart, Furse, Jones, Pechman)
Disclosure (5): Management Changes

- (These findings are) based on validation and causality audit results regarding the (ARS Persuasion) Metric as of February 2006, with updates provided to MASB in February 2008.
- Significant changes in the company’s management team occurred shortly thereafter (2006 and 2008).*
- In light of the management changes, MASB recommends an “intermittent audit” in order to ensure transparency and continuity in provider’s quality assurance programs.

Source: “Measuring and Improving the Return from TV Advertising (An Example),” MASB, April 2008

*Change in ownership occurred in 2010: The ARS Group was acquired by comScore.
Marketing Accountability Standards Board
of the Marketing Accountability Foundation

Thank you!