



## BEST PRACTICE #4

# Measurement

Overall, excellence in this best practice area is challenged by the complexities inherent in definitively measuring the impact of brand and marketing components. Excluding the private sector examples, which we did not score (measurement data are typically confidential), there was strength in two of the four sub-elements – testing message effectiveness and evaluating program results; however, the second two areas, measuring brand performance and marketing vehicle effectiveness, represent opportunities for improvement.

### TEST MESSAGE EFFECTIVENESS

Without testing marketing messages and approaches, practitioners cannot be certain whether the messages are indeed clear and persuasive, or whether the approaches have the desired effect. This seems clear, as this was the strongest best practice sub element that we evaluated. All organizations tested marketing elements before launch. The Brothers for Life TestaBoy campaign, the one example that we did not score favorably, seems to have been an anomaly, as there were other Brothers for Life campaigns that had been tested and improved before launch.

There was considerable variability in the type of testing methodology, which is appropriate given the diversity in the types of campaigns. For example, A360 Kuwa Mjanja and Future Fab, programs that relied heavily on interpersonal communications and events, used an adaptive approach, whereby the programs were consistently monitored and adjusted based on the teams' ongoing learning. MTV Shuga and Shujaaz, two edutainment programs, both relied on qualitative research to ensure comprehension and authenticity prior to launch. They both also monitored social media to identify areas of confusion or other issues, which were then addressed in social media and / or in subsequent episodes. Most of the programs that used mass media, including Trust and Truth, used traditional copy testing techniques.

**Key take-away:** The type of test and resources allocated to testing depend on the investment in marketing vehicles, with expensive, mass media campaigns requiring more rigorous testing. For testing to be effective, implementers must allocate time and resources to adjust marketing communications and programs based on learning resulting from message testing.

### EVALUATE PROGRAM RESULTS

Program evaluation is a key step in determining whether a campaign has been effective. However, statistically validated study designs can be expensive, time-consuming, or potentially inappropriate for program goals, meaning that many brands and campaigns tend to report service delivery statistics in lieu of more objective measures. Additionally, while donors rely on program evaluations to determine whether investments have been worthwhile, not every team or initiative has the necessary expertise to execute an independent evaluation. The case studies included in this report highlight the many tensions and challenges of measuring program results, including designing evaluations, determining which statistics matter most, and identifying the correct timing of evaluations given how long it takes to change behaviors.

#### Designing the evaluation:

In many cases for branded campaigns, the “gold standard” of a randomized control trial will be impractical or inappropriate, and will almost certainly add more costs. In only one of the case studies, MTV Shuga, did they use a control group with a placebo television program to compare exposure to MTV Shuga against program objectives. There are many alternative quasi-experimental designs that can be used to measure program results. For example, ORS/zinc implemented a difference-in-difference technique to measure the impact of their mass media initiative.

### Service delivery statistics are not the same as evaluations:

The majority of case studies that we reviewed reported service delivery statistics as evidence of their program results. For example, A360 reports that they exceeded their target distribution of 20,000 contraceptive methods by 250% over the project period. However, these program data do not tell us whether the campaign was the driver of this increase, or if their initial targets were too low from the beginning.

### Program exposure vs. Program evaluation:

As has been discussed elsewhere in this report, lasting behavior change can take years or even decades to take root, as is the case with the Truth campaign that originally launched in 2000. Yet most programs are designed to fit donor funding cycles of 3-5 years, meaning that, in order to receive continued funding, interventions are forced to report success within a very short time span. The result of this tension between behavior change timelines and donor funding cycles is that the short-term metrics that are chosen to measure program objectives may not reflect real behavior change.

The case studies highlight the room for improvement in the realm of program evaluation. Some key takeaways from the various examples we reviewed include:

- Donors and program implementers alike should remember that lasting behavior change takes time. Evaluations should therefore select indicators that can realistically be measured within the project period. For larger scale projects, more complex evaluation methods may be necessary. For example, NURHI had an ambitious goal of raising contraceptive use at the population level. An evaluation of this scope therefore required an extensive door-to-door survey technique that would not be feasible in projects of a larger scale and scope.
- Evaluations of media campaigns should move beyond knowledge and awareness surveys to evaluate exposure. Instead, more creative survey methods such as focus group discussions and more interactive survey designs should ask probing questions about whether and how exposure may have influenced behavior.
- Program design should include and budget for program evaluation specialists at the outset of the campaign in order to ensure that evaluations are scientifically grounded and can be implemented in tandem with program activities. Programs such as Shujaaz that attempt to evaluate their programs after project launch will be limited in the conclusions that they can draw.

We recommend MTV Shuga, NURHI, Shujaaz, ORS/zinc, and Truth as examples of program evaluations that used statistically valid methods to conclude that they had achieved their intended results. In many of these cases, the program implementer partnered with an external research organization with expertise in program evaluation. The more successful examples also set out with a pre-determined evaluation method and conducted baseline surveys before the launch of their initiatives.

**Key takeaway:** In order to prove the effectiveness of an initiative, programs should include program evaluation expertise, budget, and time in their initial designs. Donors and program implementers should remember that behavior change takes time, and evaluation designs should reflect this fact.

## MEASURE BRAND PERFORMANCE

Most programs measured brand awareness, often to assess brand or campaign exposure; however, few measured brand attributes, which is a miss. Measuring brand attributes helps implementers understand whether the brand is working as intended. For example, over time, the Trust brand lost relevance with a younger generation of potential users. As Trust worked to appeal to this new audience, it became important to monitor brand attributes, such as “is a brand for youth” or “is used by people like yourself.”

We highlight the Truth Initiative as a best-in-class example for this best practice. Based on a study of the original Truth campaign that launched in 2000, researchers had demonstrated that higher brand equity scores were associated with lower levels of smoking uptake: *“Brand equity in truth mediated the effects of campaign exposure (measured by*

confirmed recall of campaign ads) through specific targeted attitudes about tobacco industry manipulation and a desire to remain independent of tobacco from advertising on adolescent smoking uptake.”<sup>32</sup> In 2016, Evans et al developed a brand equity model for the FinishIt campaign, with the intent of understanding whether “changes in brand equity over time may predict tobacco use behavior at a later time point.”<sup>33</sup> The study includes 18 questions that help assess four overall constructs (brand loyalty, leadership/popularity, brand personality, and brand awareness) and were added to waves 3 and 4 of the Truth longitudinal cohort study. Study results showed that FinishIt brand equity “predicts increased anti-tobacco attitudes and reduced tobacco use behavior over time (at six months).”<sup>34</sup> In addition, implementers can use the study results to understand how they can tweak marketing programs. For example, if – hypothetically – brand loyalty were to fall over time, the program team could diagnose the marketing elements to determine what might have contributed to the decline, and how best to restore it.

We recognize that most programs do not have resources or systems in place to conduct such a study; nonetheless, practitioners should add questions to evaluation studies to confirm whether the brand is working in the way it was intended. For example, the ORS/Zinc project included six questions to understand attitudes about ORS and zinc that the campaign was intended to address (e.g., “ORS and zinc is recommended by health providers,” “ORS and zinc is more effective than home remedies”).

Several of the projects we assessed (e.g., NURHI, Brothers for Life, Shujaaz) were grounded in learning that the audience had felt judged by other “finger wagging” programs. Several of the questions that Shujaaz includes in audience surveys, e.g., “I have a positive image of Shujaaz,” “Shujaaz has a positive impact on my life” and “I trust Shujaaz,” can help implementers understand whether the brand is a positive source of inspiration and encouragement, a finger-wagging health program, or somewhere in between. We also like the question, “I have recommended Shujaaz to other people.” As Frederick Reichheld described in the HBR, “When customers act as a reference, they do more than indicate that they’ve received good economic value from a company; they put their own reputation on the line. And they will risk their reputations only if they feel intense loyalty.”<sup>35</sup>

**Key take-away:** Project teams do a good job of assessing awareness but need to routinely measure brand associations. In the absence of a robust brand equity model, implementers should develop questions to assess whether the audience perceives the brand the way it is intended.

## EVALUATE MARKETING VEHICLE EFFECTIVENESS

Very few programs evaluated marketing vehicle effectiveness, likely due to the limited number of marketing vehicles used by many of the programs in our case study series, as well as challenges of collecting accurate data. Those that did include such an evaluation relied on self-reported data, which has obvious limitations, but is also the most practical approach. Some programs, such as Brothers for Life, did a detailed evaluation of a specific vehicle (e.g., Facebook), to understand how well the videos shown on the social media site – as well as the resulting conversations – affected audience behavior. What is lacking is a comprehensive evaluation that assesses the relative importance of each vehicle in terms of reach, impact, and cost.

**Key take-away:** In the absence of robust data that prove audience viewership and link exposure to behavior, measuring marketing vehicle effectiveness remains a challenge. Project teams should seek to understand not only where the audience may have been exposed to a particular message, but the degree to which each vehicle influenced behavior.

---

32) Evans, et al, “Effects of the truth finishit campaign on tobacco outcomes,” *Prev Med Rep*, Dec 2, 2017

33) *Ibid*.

34) *Ibid*.

35) Reichheld, Frederick, “The One Number You Need to Grow,” *Harvard Business Review*, December 2003.