Addressing Challenges of Testing Large Numbers of Messages—Effective Modeling of Interactions

Coping with the increasing number of test messages and requirement for quantitative testing

Increased pressure to demonstrate value for money in marketing campaigns has had a knock-on effect on market research, which must now deliver more critical assessments, leading to an increasing demand for quantitative message testing. The consequences of this are that: (1) there is a requirement to quantify message effectiveness; (2) many more messages are now tested within a research exercise. This increase has been fueled both by volume and diversity: a greater number of messages and supporting statements to test; and multiple messages aimed at specific segments or at addressing specific points. The objective of this effort is to deliver more complex messages to differentiate agents in increasingly crowded categories. Measuring interactions is critical in building effective message strategies and individual messages. However, the number of interactions between messages increases rapidly to the point where they cannot all be measured reliably in a survey, for example 20 messages would yield 190 two-way interactions (or 1,410 three-way interactions, or 15,504 five-way interactions).

What are interactions?

An interaction (positive or negative) occurs when the impact of two or more messages together is more or less than the sum of the impact of the individual messages. To take a consumer example, consider the preference for chocolate and nuts. While someone may like both chocolate and nuts, they could dislike having chocolate and nuts combined, in a candy bar for instance. Thus, neither condition alone will result in a dislike; only in combination do they have this effect.

Considering Interactions Between Messages

For the above reasons, it follows that approaches that served us in the past no longer provide the rigor needed. Traditional techniques simply measure the value of individual items (proportion of respondents rating or ranking items). Other approaches used in advertising such as TURF (Total Unduplicated Reach and Frequency) and BUNDOPT (Bundle Optimization) are flawed in the context of message testing, mainly because they fail to take into account interactions. For different reasons, traditional conjoint techniques are inadequate because they cannot deal with the potential number of interactions highlighted above. Three common examples of message interactions follow.

- The top rated messages may all make the same point (be in the same 'category', e.g., efficacy)—minor wording changes usually result in similarly rated messages that add little in combination.
- By contrast, some combinations of messages may reinforce each other, providing more impact together than alone, yet this synergy is not uncovered by traditional techniques.
- Some messages may not rate highly in isolation, but their inclusion in the *'message bundle'* may be critical to provide credibility and impact. A 'cost of entry message', for instance confirmation that side effects are equal to current products, may have to be included in the message bundle, even though it has little impact by itself.

Our Advanced Analytics group has researched this need: an advanced approach to identify, measure and resolve these interactions.

An Advanced Approach that Resolves the Interaction Problem

Concept—getting to the heart of the problem

It seemed that the interaction problem needed to be handled at the category level, where the number of

Addressing Challenges of Testing Large Numbers of Messages—Effective Modeling of Interactions (continued)

potential interactions was more manageable. We also needed to measure the impact of additional messages within discrete categories. For example, within the efficacy category, it is important to consider the primary efficacy message(s), as well as the effect (or not) of any secondary efficacy message(s) such as time to onset or duration of action, or the effect of including outcomes measures. Again, the issue is whether the addition of messages within a category adds more or less than would be expected by simply adding the values of each message.

Methodology and Modeling

This interactions modeling approach is based around a conjoint method which measures interactions between 'message sets' to determine optimal message bundles. Message sets are generated via an experimental design. The method predicts the relative impact of each message and message bundles on a dependent variable, such as likelihood to prescribe. It can test up to 50 messages (30 messages are optimal) and provides output in an Excel-based user-friendly simulator.

A two-stage approach is used.

- First Stage: individual messages are classified into categories (e.g., mode of action, efficacy, other categories) either
  - *A priori* (through early research or intuition) or
  - *Post-hoc* factor analysis
- Second Stage: a multiple-step process is used to determine the optimal message bundle
  - Step 1: Individual message assessment
  - Step 2: Assessment of message set combinations
  - Step 3: Composite impact analysis.

In Step 1 of the Second Stage, messages are individually rated on ability to impact prescribing relative to all items generally, and relative to other items in the same category. In Step 2, individual messages are grouped by the experimental design computer program, and these message sets are rated based on ability to impact prescribing for each patient type, using a choice exercise. In Step 3, the relative impact of each statement on prescribing intentions is assessed. A message simulator allows the impact of different combinations of messages to be tested. There may be diminishing returns as the number of statements for each category increases. Modeling helps build potential message strategies by revealing the implications by specialty, by segment, and by patient type. This methodology fits well into either face-to-face or Internet research techniques.

Issues to Consider

In preparing for a quantitative message test there are several points we consider. We ensure there are substantive differences among the messages to be tested in a particular category. These differences are determined through qualitative research, although judgment calls may need to be made if similar messages show similar results in the qualitative phase. We try to limit the number of possible test messages to a maximum of 30 overall, and 5 or 6 within each category. We research which categories are critical to the product's story (e.g., efficacy: time to onset), and which are less important. Where possible, we include concise, easy-to-read graphics/tables to support/ evidence the message.

Typical Outputs and Benefits

Modeling shows the impact of including multiple items from the same category in the composite message. Modeling also reveals the impact of the number of categories included in the composite message set.

The simulator provides insight into wider marketing strategies for example:

- Modeling can show the optimum message bundles for a specialist versus primary care audience.
- The simulator can also show the impact of shorter message bundles versus longer message bundles, which can provide useful information on the best messages to bundle in situations where a representative only has time for a short detail.
- Repeated over time, research can also show changes in perception of messages (for instance, pre-launch ➔ peri-launch ➔ post launch). This helps clarify those messages that have become accepted and no longer need communicating in great depth, as opposed to those messages that need greater focus to help differentiate the brand.

Our conceptual framework is that messages work both at the individual level and also through their interactions with each other. Therefore measuring these interactions is critical in building effective message strategies and individual messages.

Conclusions

An advanced interactions modeling approach measures and resolves the interactions that are critical in driving effective message strategies and determining optimal message sets of different sizes. Essentially this composite impact analysis methodology is able to handle the very large numbers of messages and supporting statements that need to be tested to deliver complex messages, aimed at specific segments, which are required to differentiate agents in increasingly crowded categories. The methodology generates assessments of messages as individual components and in different combinations which are then grouped into message sets. The approach incorporates a conjoint-like exercise but importantly, also includes modeling techniques designed to overcome the limitations of simple conjoint-like exercises (i.e., their inability to represent the 'interactions' that occur between individual categories). This approach (1) measures 'interactions' between messages/sets to determine optimal message bundles, and (2) predicts relative impact of each message, and message bundles, on likelihood to prescribe. It can test up to 50 messages, 30 is optimal, and provides output in an Excel-based user-friendly simulator. Variations of this technique can be applied to test a wide variety of outputs other than messages, such as concepts, positioning statements, advertisements, and visuals.

Larry Goldberger,
Advanced Analytics, Adelphi Research by Design
www.adelphigroup.com/marketresearch