

# Concept Testing New Products: Errors vs. Segmenting Variables

Abram Poczter, Ph.D., New York Institute of Technology, USA

## ABSTRACT

*Utilizing data from an actual concepts testing experiment, this paper presents a statistical analysis of the impact of problem- and benefit-oriented concepts statements as well as several “variables-in-audience” such as: Product Involvement, Use of the Product Category, Level of Compliance, and Level of Education. These predictors are evaluated in terms of their impact on the criterion variables, operationalized as components of an emerging attitude, and are discussed as either the elements of the segmentation scheme, or as errors biasing the representativeness of the sample.*

**Keywords:** Concept Testing; Cognitive-Affective-Conative Components of Attitude; Hierarchy of Effects; CAD Personality Inventory; Product Involvement

## INTRODUCTION

Concept Development and Testing occupies a precarious, yet important position in the process of new product development. Upstream, there is the Fuzzy Front End (Smith and Reinertsen, 1992), where the demands of business/marketing strategy, R&D and technological capabilities, and external environmental trends interact in giving birth to a new product idea (Ideation stage). Downstream from the Concept Testing, there are stages that consume exponentially increasing amounts of time and money, such as Prototype Development and Testing, Pilot Production, Market Testing, Test Marketing, and Roll-out. In addition to the “out-of-pocket” expenses, there are opportunity costs resulting from the exposure of a product as well as the marketing plan to competitors. This is well understood by both practitioners and academic researchers.. In a recent, broadly based survey, a vast majority of companies reported utilization of this stage in situations ranging from existing product improvement to radically new products (Peng and Finn, 2008).

Although practitioners are typically unwilling to reveal specific techniques and methodologies used as well as their efficacy, there is a clear tendency to treat Concept Development as a process of bridging the mind of the architects of the Fuzzy Front End and the minds of the consumers. Concept Testing, in turn, is treated as the test of the ability of the NP concept to generate positive attitudes, including positive behavioral intentions.

## Elements of Concept Development and Testing

A simplified progression in concept development and testing is presented below:

1. Strategic Business considerations call for a search for new product ideas in specific strategic business clusters.
2. New product ideas are generated for specific (existing or new) market segments.
3. New product ideas are turned into concept statements (or stimuli), varying from a simple description of functional characteristics, to promotional copy, to a rough rendition of a printed advertisement or a story board, to a “mock-up” or a prototype.
4. A measuring instrument, usually a questionnaire, is developed, to capture respondents’ attitudinal and conative reactions to the concept (hierarchy of effects). In addition, the questionnaire ought to contain questions about respondents’ demographic backgrounds, personalities and behaviors.
5. The samples of respondents are exposed to the concept statement (stimulus) and their responses are captured via the measuring instrument.

6. The responses are evaluated in terms of the ability of the concept stimulus to convey knowledge of product benefits, image of the new product, liking, and intention to try/buy the new product.

### **Statement of the Problem and Literature Review**

In the taxonomy of research designs, Concept Testing belongs to the causal research category, relying on experimental designs to test the effects of exposure to a concept stimulus.

Every step in concept development and testing is yet another chance to fail to control for factors jeopardizing validity and reliability. Following Klink and Athaide (2006), the fundamental sources of concept-test design errors stem from:

1. Concept statement (stimulus) design  
Examples would include unclear statements of product characteristics, emotional appeal, when consumers base their decision on reasoning (and vice-versa), etc. This tends to effect the internal validity of the experiment.
2. Response measurement  
Does the instrument measure what it is supposed to measure? This speaks to construct and content validity.
3. Respondents' selection  
Do the respondents' profiles match the target market segment's? Do the personality and psychographic characteristics confound the responses to the concept?
4. The compression of adoption and diffusion processes into a brief, discreet-time reaction
5. Network externalities, important in the case of new product concepts representing discontinuous innovations.

In this paper, the emphases are on areas 1 through 3.

**Design of the concept statement** rests on the principles and methodology of the communicational research. Rather than attempting development of a general theory of message effect, the approach is one of empiricism. Every instance of research is a test of the efficacy of some type of message structures deemed appropriate in a given situation, such as; One- vs. Two-sided, Low vs. High Level of Fear, Climax vs. Anti-climax, etc.

In this case, the new product was a hair coloring product aimed at the "in-home" segment of non-users. Based on the definition of an attitude toward an object reflecting a person's perception of the object to provide a benefit or eliminate a problem (see p.3), it was decided to develop two different concept statements. A problem-oriented statement emphasized the new product's efficacy in avoiding problems usually associated with hair coloring, such as unnatural shade, uneven fading, dull colors and damaged and unmanageable hair. A benefit-oriented statement described the benefits of a natural shade, long lasting, bright natural colors, and healthy, manageable hair. In other words, the attributes addressed in both statements concerned the same product characteristics and the copy layout was identical. Following Crawford and Di Benedetto (2003), both statements were clear and realistic, which was corroborated in a pretest.

**Response measurement** was accomplished via an extensive questionnaire, consisting of items measuring:

- Perceived Instrumentalities
- Level of development of an attitude
- Level of involvement with the product category
- Use of the product category
- Personality typology

### **Level of Perceived Instrumentalities**

Perceived Instrumentalities (Beliefs-in) mentioned in the concepts statements were measured on 5 point, Likert-scale like items. These items measured the strength of informational beliefs developed because of exposure to

the statements. Attitude toward an object, concept, or idea is a function of the perception of the instrumentality of the object, concept, or idea to bring about a positive state (benefit), and/or block a negative state (eliminate a problem). The type of benefits or problem solutions were of both a functional nature, such as hair manageability, covering gray hair, natural shade, and socio-psychological in nature, such as being satisfied with oneself, receiving positive comments from others, etc. This is consistent with multi-attribute models and the instrumental school of attitudes (Rosenberg, 1956, 1960; Fishbein and Ajzen, 1975).

Similarly, levels of Perceived Instrumentalities (Beliefs-in) not mentioned in the concepts statements were measured on 5 point Likert-scale like items. These items measured the strength of inferential beliefs developed because of exposure to the statements. These consumer beliefs “round-up” the image of the product. Questions such as “What do you think the price of the products is going to be, as compared to competition?” “What other products would you consider as complementary?” “What type of stores would you expect to find this product?” provided the ultimate in customer orientation, by allowing consumers to design elements of the marketing mix.

### **Hierarchy of Effects**

Attitude’s tripartite structure (Cognitive-Affective-Conative) when used in a dynamic sense of attitude development, leads to one of the following three basic paths (Ray, 1973):

- High Involvement (Cognitive→Affective→ Conative) or (Learn→Feel→Do)
- Low Involvement (Cognitive→Conative→Affective) or (Learn→Do→Feel)
- Dissonance Attribution (Conative →Affective→Cognitive) or (Do→Feel→Learn)

The path selected depends on the individual’s level of involvement with the product category, and not the product classification in terms of cost, durability, or tangibility, although there is undeniable correlation (Belch and Belch, 2007). Thus, it is possible that exposed to highly attractive concepts, some consumers would express willingness to buy, without expressing affect (low involvement segment), and some would follow the more traditional path of expressing affect followed by the willingness to buy. Thus, the Hierarchy of Effects model is useful in interpretation of the results, i.e. evaluation of the concept.

The Dissonance-Attribution hierarchy, requiring actual behavior, was not considered in this study.

- The measuring instrument “followed” the respondents along the Cognitive→Affective→ Conative components of the hierarchy of effects.
- The questions about perceived instrumentalities measured the extent to which the Cognitive component has developed.
- Question regarding overall liking and the affective affiliation, measured the development of the Affective component.
- Questions related to the likelihood of coloring hair, and using the new brand in the process, were used to measure the development of the Conative component.

### **Involvement**

Low-High involvement dichotomy emerged in communication research with the publication of McLuhan’s “Understanding Media (1964)”, and “The medium is the message: an inventory of effects” (with Quentin Fiore, 1967). In addition to heralding the arrival of an “electronic village,” television was thought to be a “cold” or “low involvement” medium, for it required minimal involvement of the viewer for information processing. Krugman referred to the effects of television as “memory without recall, learning without involvement” (Krugman, 1965, 1971). Therefore, a Readership of Print Advertisements question (with Yes or No as response categories) was used as operationalization of involvement. “Yes” classified the respondents as a “high involvement” type.

**Respondents’ selection** was reflected in measuring their various characteristics.. Per se, it was not so much a selection, as there were no a priori target segments. The study, utilizing a national, representative sample attempted to ascertain the impact of consumer descriptors, i.e. would uncover groupings of consumers varying in terms of their responses to concept stimuli.

Personality characteristics, by definition, mediate a person's responses to the external environment. The problem of utilizing personality measuring instruments is that psychologists calibrated their instruments towards responses to the social environment, rather than marketing stimuli. Thus, after having selected an instrument with high facial validity, a marketing researcher needs to validate it, within the marketing context.

In this research, CAD (Compliant-Aggressive-Detached) Scale (Cohen, 1967), an operationalization of Horney (1947) personality theory, was utilized. Compliance, Aggression and Detachment typology seems to have relevance to the concept testing situation. A compliant type is most eager to please, comply with group norms – in other words, a happy follower. In a concept testing setting, it is hypothesized that compliant individuals will tend to “pull back” with the conative component of the attitude. Thus, compliant individuals would tend to buy a new product only after others bought it. Therefore, if the group of participants in a concepts test is dominated by compliant individuals, the “Intention to buy” question will be a poor predictor of the attractiveness of the concept, i.e. it would underestimate the proportion of likely buyers.

Aggressive individuals who tend to strive for leadership of a group develop and enforce values and rules. They will tend to be early new product adaptors, so they can influence the followers (Compliant) to follow suit. Thus, they would tend to bias upward the Intention to buy question.

Detached individuals move away from people, do not aspire to group membership and/or group leadership. Their responses to a concept statement are based on their own evaluation. The CAD Scale utilized 15 questions measuring the level of aggressiveness, and used 10 questions each to assess Compliance and Detachment.

**Present Use of the Product Category** was measured because of its universal utilization as a segmenting variable in new product marketing. New products are targeting either present users (light- or heavy-half) or non-users. A concepts test of respondents as a group, must reflect the targeted segment; otherwise, the results do not have external and predictive validity.

Moreover, users (of product category) were shown to differ from non-users, in their responses to advertising (Wheatley and Oshikawa, 1970; Wheatley, 1971). Finally, users would tend to have a higher level of knowledge of the product category, which mediates concept evaluation (Moreau, Lehmann and Markman, 2001).

Level of Education of the respondents (Less than HS, HS/Some college, College+) was included as it seems to be an essential component (used implicitly or explicitly) in copy strategy. Concept statement development and concept testing bear close similarity to copy development and copy testing. Moreover, Lehmann (1977) demonstrated a close link between the levels of education and the pattern of responses along the hierarchy of effects.

## **RESEARCH DESIGN**

A sample of 1000 women was drawn from a National Consumer Panel, managed by Home Testing Institute (a subsidiary of NPD Group, NY). The sample is demographically representative of the US population of women residing in two or more person households. The sample was randomly split into three groups; two experimental and one control. With n=1000, randomization ought to eliminate selection bias.

Experimental groups received by mail, either the Problem or the Benefit oriented message, and the questionnaire, that included the attitudinal responses along the hierarchy of effects, and respondents' characteristics. In the Control group's questionnaire, the questions about the specific new brand of hair coloring were replaced with attitudinal questions about advertising of hair coloring products in general. Thus, the research design closely followed After-Only with Control Group design, known to control for most of the factors potentially jeopardizing external and internal validity (Churchill, 1995). The response rate was 62.3%, evenly distributed across the three groups, indicating a limited non-response error, and an absence of reactive arrangements.

## **ANALYSIS**

### **Stage 1 - Results and Discussion**

All three groups were analyzed in terms of the respondents' background characteristics and the attitudinal responses. Essentially, it was a test of whether the groups were equivalent and whether the exposure to the concept statement, i.e., experimental treatment, had a significant communicational effect. As expected, considering the sample size and the random group assignment, the three groups did not differ in terms of background characteristics, such as demographics and use of the product category. Statistically significant differences between the three groups were found in the cases of Product Category-specific Confidence and the Perceived Media Informativeness. Thus, the above are not descriptor variables-in-audience, but rather dependent variables affected by the treatment. Therefore, these two variables were removed from the set of independent variables. All in all, these analyses indicated that the experiment is free of selection bias.

#### **Cognitive Component - Perceived Instrumentalities and Perceived Comprehension**

As noted, there were two types of perceived instrumentalities, functional and socio-psychological, with a very different pattern of results.

With the functional perceived instrumentalities, a Duncan Multiple Range Test indicated two subsets, one consisting of the two experimental groups, and one consisting of the control group. The exceptions, where Duncan's Multiple Ranges Test has indicated three homogeneous subsets were:

- Question about an attribute stated in a problem-oriented fashion (Can avoid an uneven fading). The problem-oriented concepts statement was more believable, (as compared to the Benefit-oriented statement and the general hair color products advertising) when addressing just one attribute only, uneven fading.
- Two questions aimed at checking whether the problem- and benefit-oriented statements were indeed perceived as such (VAR010 and VAR011). It is a very welcomed result, as it implies a high degree of construct validity.

With the socio-psychological perceived instrumentalities, Duncan Multiple Range Test indicated an existence of a single homogeneous set comprised of all groups, implying there was no difference between the experimental and control groups, thus the concept test did not have a statistically significant effect.

#### **Effective Component**

The difference between the experimental and the control groups, in terms of Overall Liking was not statistically significant. Thus, the effective component did not follow the development of a cognitive component. Perhaps it is related to the finding that the socio-psychological perceived instrumentalities did not differ across the groups.

#### **Conative Component**

The experimental groups, on average, scored significantly higher than the control group in terms of the intention to try/buy the hair color brand. Thus, the emerging pattern is of a Low-involvement type, Development of a Cognitive component is sufficient for the respondents to develop the Conative component, an intention to try/buy the tested product.

### **Stage 2**

For further analysis only the Experimental groups were retained, for a clearer contrast between the Problem and Benefit-oriented statements.

The final ANOVA model consisted of the Criterion variables (Knowledge, Liking, and Intention-to-Buy) in turn serving as dependent variables, and the following set of Predictors (Independent Variables):

- Type of the concept statement (Problem, Benefit)
- Level of Involvement (High, Low)
- Level of Compliance (Low, Medium, High)
- Use of the Product Category (Yes, No)
- Level of Education (<HS, HS/Some college, Coll+)

In order to avoid the problem of inter-correlation between the criterions variables on each consecutive model's significance level, a "Step-down F" procedure was utilized (Stevens, 1972).

Essentially, each successive model utilizes as a covariate the independent variable form the preceding model. Thus, analysis of variance of Liking was performed with Knowledge as a covariate, and in case of Intention-to-Buy, Liking was used as a covariate.

## **Stage 2 - Results and Discussion**

**Message effect – The Type of Concept Statement (Benefit- vs. Problem-oriented) had no significant differential effect, at any of the stages of attitude development.**

### **Cognitive Component – Perceived Comprehension**

Multiple Classification Analysis (MCA) revealed that Involvement and Compliance were positively related to Perceived Comprehension. The relatedness of Education was not "linear." Respondents on either end of the education spectrum had group means below the grand mean on Perceived Comprehension.

Use of the Product Category was not significant, i.e. holding all other factors equal, whether the respondents colored their hair or not was not seen to impact the level of comprehension of the concepts statements.

### **Affective Component – Overall Liking**

Main effects of Involvement, Level of Compliance, Use of the Product Category, and Education were significant beyond 0.06. MCA revealed a positive relationship between Overall Liking of the concepts tested brand and Involvement, Level of Compliance, and Use of the Product Category. Once more, the relationship of Education and the criterion variable was not "linear."

### **Conative Component – Intention to Buy (Brand)**

Main effects of Involvement, Use of the Product Category, and Education were significant beyond 0.05. MCA revealing a positive relationship between Intention to Buy (Brand) and Involvement and Use of the Product Category. Once more, the relationship of Education and the criterion variable was not "linear." Level of Compliance was not significantly related to the intention to buy. This is expected, as the compliant individuals are followers, rather than early adaptors. They buy new products only after the new product becomes a norm of the group.

## **CONCLUSIONS**

Concept testing is by its nature an experiment, where a group(s) of participants is exposed to a new product concept, and intensity of incipient attitude is measured. Considering time and money commitments, this stage consists of a series of experiments with an increasing level of resources allocated to the level of execution of the concept, sample size and quality, and the experimental arrangements. For example, as an initial stage, a "one shot case study" may be undertaken where the concepts statement is just a functional description of the product and the sample is a convenience sample of the firm's employees. This arrangement, lacking rudimentary controls over factors jeopardizing internal and external validity cannot be used for anything but eliminating concepts, reworking them, and re-testing with the same sample.

Eventually, the concept testing experiment needs to live-up to its scientific roots.

- A. The concept statement ought to contain the level of creative execution appropriate for the degree of newness of the product. New product ideas, classified as discontinuous innovation, ought to have concept statements that are simple descriptions of product characteristics. New product ideas that are continuous innovations ought to contain the level of creative execution compatible to the brands in the market, for the respondents will be comparing new products statements to those of competitors. In this research, the concept statement was in a form of a “rough” of a printed ad, since print advertising was planned to be the “workhorse” of the advertising campaign. Since there was no differential impact of message type on the development of an attitude, management has used a mixture of problem- and benefit-oriented claims in the introductory campaign.
- B. The sample, in terms of size, selection and composition (variables-in-audience) must assure an acceptable level of the Standard Error of Estimate, as well as a high level of external validity, i.e. ability to generalize the results to the population of interest. This means that a particular descriptor may represent either a bias or an error in concept testing, Ling and Finn (2008), or a segmenting variable. For example, let us consider Level of Compliance. The positive relationship (above average) of this variable with the Perceived Comprehension and Overall Liking did not translate into above average intention to buy (conative component). Furthermore, assume that highly compliant individuals are over-represented in the sample. Whether it is an error depends on which segment is targeted in the initial campaign. Indeed, if the target consists of aggressive people, the sales forecast based on intention to buy will be biased downwards.

#### **ACKNOWLEDGEMENT**

The author would like to acknowledge the invaluable inputs of Dr. Spencer F. Tinkham, throughout the project.

#### **AUTHOR INFORMATION**

**Dr. Abram Poczter** is a professor in the Department of Marketing Sciences, School of Management, New York Institute of Technology. Teaching and research interests revolve around Marketing (Consumer Behavior, Customer Value and Satisfaction, New Products Management, Marketing Research, International Marketing, and Corporate Strategy). E-mail: [apoczter@nyit.edu](mailto:apoczter@nyit.edu)

#### **REFERENCES**

1. Churchill, Gilbert, A. (1995). *Marketing Research: Methodological Foundations*. The Dryden Press.
2. Cohen, Joel (1967). An Interpersonal Orientation to the study of Consumer Behavior. *Journal of Marketing Research*. vol. 4.
3. Crawford, Merle and Anthony Di Benedetto (2003). *New Products Management 7/e* McGraw Hill International.
4. Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*. Reading, MA: Addison-Wesley.
5. Horney, Karen (1947). *Our inner Conflict*. Morton: New York.
6. Klink, Richard R and Gerard A Athaide. (2006). An Illustration of Potential Sources of Concept Test Error. *The Journal of Product Innovation Management*, 23(4), 359-370.
7. Krugman, Herbert (1965). The Impact of Television Advertising: Learning without Involvement” *Public Opinion Quarterly*, Vol. XXIX.
8. Krugman, Herbert (1977). “Memory without Recall, Exposure without Perception” *Journal of Advertising Research*, Vol. 17.
9. Lehmann, Donald, R., (1977). Response to Advertising a New Car. *Journal of Advertising Research*, vol. 17, No 6.
10. Ling Peng, Adam Finn. (2008). Concept testing: the state of contemporary practice. *Marketing Intelligence & Planning*, 26(6), 649-674.
11. McLuhan, Marshal (1964) *Understanding media: the extensions of man*. Cambridge, MA: MIT Press.

12. McLuhan, Marshal and Quentin Fiore, Co-ordinated by Jerome Agel. (1967) *The Medium is the Massage*. Bantam Books / Random House.
13. Moreau, Page, C., Lehmann, Donald, R and Arthur B. Markman (2001). Entrenched Product Knowledge and Consumer Response to New Products. *Journal of Marketing Research* 38: 14-29 (February).
14. Rosenberg, Milton, J. (1956). Cognitive Structure and Attitudinal Effect. *Journal of Abnormal and Social Psychology*, 53 American Psychological Association.
15. Smith, Preston G., Reinertsen, Donald G. (1992). Shortening the Product Development Cycle. *Research Technology Management*, 35(3), 44.
16. Stevens, J.D. (1972). Step-down Analysis and Simultaneous Confidence Intervals in MANOVA. *Multivariate Behavioral Research*, vol. 8.
17. Wheatley, John J., and Sadaomi Oshikawa. (1970). The relationship between anxiety and positive and negative advertising appeals. *Journal of Marketing Research*, 85.
18. Wheatley, John J. (1971). Marketing and the Use of Fear- or Anxiety-arousing Appeals. *Journal of Marketing* 35, 62.