eISBN: 978-1-60805-284-4
ISBN: 978-1-60805-419-0

Rule Developing Experimentation
A Systematic Approach to Understand & Engineer the Consumer Mind

Editors:
Alex Gofman and Howard R. Moskowitz
Moskowitz Jacobs Inc.
USA

Bentham e Books

Rule Developing Experimentation: A Systematic Approach to Understand & Engineer the Consumer Mind

Edited By

Alex Gofman, Ph.D. (Deceased)

&

Howard R. Moskowitz, Ph.D.

Moskowitz Jacobs Inc.,
White Plains
New York
USA

DEDICATIONS

To the memory of my co-editor, Alex Gofman, who in his short 52 years accomplished many things. Alex came here with the burning hope of many immigrants to the United States. From the start, Alex was filled with the dreams of what he could accomplish, and turned many of them into realities. His efforts eventuated in a wonderful family, new technology that would help people and companies, and the continuation of his education, culminating in the Ph.D. Alex was a sui generis. He will be missed.

Howard R. Moskowitz

CONTENTS

Foreword *i*

Preface *iv*

List of Contributors *x*

CHAPTERS

PART I: THEORETICAL FOUNDATION OF RULE DEVELOPING EXPERIMENTATION

1. **Origins of RDE and the Role of Experimentation in Consumer-Driven Innovation** **3**
   A. Gofman

2. **Isomorphic Permuted Experimental Designs in Conjoint Analysis** **34**
   A. Gofman and H.R. Moskowitz

3. **Detecting Explicit and Implicit Interactions within Rule Developing Experimentation** **48**
   A. Gofman and H.R. Moskowitz

4. **Putting RDE on the R&D Map: A Survey of Approaches to Consumer-Driven New Product Development** **72**
   A. Gofman

5. **RDE in Concept Research: An Empirical Demonstration** **90**
   A. Gofman and T. Mets

6. **Consumer Metric Scales** **106**
   K. Lopetcharat and D. Paredes

7. **Foundation of Sensory Optimization in the Food Industry** **129**
   H.R. Moskowitz

PART II: PRACTICAL SENSORY OPTIMIZATION

8. **Introduction to Sensory Optimization** 145
H.R. Moskowitz

9. **Sensory Optimization in Research and Development** 158
M. Gacula Jr.

10. **Measuring Interest and Price for Sensory Experience: Application to Hotels** 184
J. Fyrbjörk, R. Shofu, L. Ettinger Lieberman and H.R. Moskowitz

PART III: MESSAGE OPTIMIZATION

11. **Messaging Across National Markets–Effectiveness and Segmentation** 201
G. Gabay, S. Rabino and H.R. Moskowitz

12. **Extending RDE to Evaluating Potential Social Anxiety Factors** 221
H.R. Moskowitz, A. Gofman, G. Gabay, H. Ashman and I. Ray

PART IV: ADVERTISING RESEARCH AND BRAND COMMUNICATIONS

13. **Consumer-Driven Advertising Research** 242
J. Ewald

14. **The ROI of Woo: Starting, Sustaining and Improving the Relationships Business and Brands Have with Consumers** 263
M. Gadd and J. Gadd

PART V: STRUCTURED PACKAGE AND WEBSITE OPTIMIZATION

15. **Helping Packages Get Noticed on the Shelf Using RDE** 279
A. Gofman, H.R. Moskowitz and T. Mets

16. **Introduction to Consumer-Driven Optimization of Landing Pages** **299**
A. Gofman

17. **Consumer Driven Website Optimization** **310**
A. Gofman, H.R. Moskowitz and T. Mets

18. **Consumer-Driven Website Customization: The Need to Manage Costs and Benefits** **327**
E.K.R.E. Huizingh and M.L. Teerling

PART VI: MIND GENOMICS®: HOLISTIC UNDERSTANDING OF CONSUMERS

19. **Introduction to Mind Genomics®** **352**
A. Gofman

20. **Mind Genomics®: A Systematic Consumer Research** **355**
H.R. Moskowitz, A. Gofman, J. Beckley and H. Ashman

PART VII: EXPERIMENTATION AND CONSUMER-DRIVEN INNOVATION

21. **Consumer-Driven Innovation Management** **385**
A.D. Narasimhalu and S. Mitra

22. **Neuromarketing 2.0: How Rule Development Experimentation is Innovating Neuromarketing Research** **405**
S. Zace and H.R. Moskowitz

23. **Balancing People's Future Demand and Design Genius** **434**
M. Bevolo

End Note **448**

FOREWORD

This book offers an intellectual delight and practical value for three audiences: (a) decision makers in a variety of fields ranging from corporate executives to public policy decision makers, (b) consumer and marketing researchers, and (c) that wonderful of all gifts, an informed reader who is intrigued by surprising insights and new scientific paradigms.

The basic premise of the book is that to address today's challenges, one *must* understand the mind of the consumer and other relevant stakeholders. In this respect, the book follows the rich tradition of marketing orientation and its focus on understanding the consumer. The book furthers this approach by advocating and illustrating the value of *Rule Developing Experimentation* (RDE), which Howard Moskowitz and his colleagues have been developing over the last 30 years, and which was introduced in *Selling Blue Elephants: How to Make Great Products that People Want Before They Even Know They Want Them* (2007, Wharton School Publishing). This new book advances THE foundational ideas by offering for each audience segment a number of relevant chapters rich in concepts, innovative approaches, insightful findings, and timely examples.

Whatever your reading style – whether you devour books cover-to-cover or skim selected chapters – this book is a "must-read":

1. The book makes a strong case for the value of truly understanding the mind of the consumer in the solution of complex managerial challenges. The challenges span the wide range of issues encountered by today'sbusiness, including R&D and new product development, sensory optimization for food products, packaging design, pricing, advertising, website customization and optimization, segmentation, as well as other key business-relevant decisions.

2. The book provides a thorough explanation of the Rule Developing Experimentation (RDE) approach, including its origin, its intellectual and compuational relation to conjoint analysis and other powerful analytic method. Additionally, the book presents a variety of creative

applications Which illustrate RDE in action, and delightfully engage the reader with innovative out-of-the-box solutions to key challenges facing companies, societies and individuals. Importantly, there is the ring of practicality, of experience, of stories and theory. The discussions are presented from the point of view of the practitioners, who conducted and wrote most of the chapters.

3. The book gives the reader a new perspective on the emerging scientific paradigm of Mind Genomics®. Mind Genomics® maps the consumer's dimensions of experience, creating microsciences of the everyday, a radically new vision for the project of consumer science. For any domain of human life, Mind Genomics® turns the spotlight on, using its RDE tool to identify the phrases which constitute for the respondents the domain of investigation , and within it what is important and what is not. The science focuses on the consumers at large, and reveals new-to-the-world, often quite fascinating, mindset segments. Whereas conjoint analysis and other approaches have done this before for specific applications, the unique feature of the Mind Genomics® world view is the relentess focus on cumulative knowledge across applications, with the goal to develop a usable, generalizable, accessible database of the findings. This ambitious goal finds its early application in this volume, but encourages the readers to add their own applications. This benefit is so important that an alternative title for this intriguing book could have been "The New Science of Mind Genomics®," or "Mapping the Consumer Mind" to highlight this breakthrough idea of the book. Whereas the additional concepts embedded in the current title are important and the book delivers on them in a thorough and engaging way, the truly innovative idea presented by Howard Moskowitz and his colleagues in their study of Mind Genomics is a bonus for readers interested in exploring the practical value of a developing scientific trend.

The three benefits of the book are delivered in engaging and insightful ways, WAYS that I hope will stimulate readers to refocus their attention on the understanding of the consumer's mind. As a professor I dearly hope that the reader will be sufficiently inspired, or perhaps simply intrigued enough to experiment with the RDE and related approaches. In the end, if the vision of mind genomics® is destined to come to fruition, in whatever format. Perhaps the readers will be inspired to augment whatever study they conduct with a search for empirical generalizations, and join the grand project of developing and continually updating the inventory of Mind Genomics® data, and its gift of insight.

Yoram (Jerry) Wind, Ph.D.

The Lauder Professor and Professor of Marketing
The Wharton School, University of Pennsylvania
Philadelphia
Pennsylvania

PREFACE

Mind Genomics®, Rule Developing Experimentation (RDE) and a New Science of the Everyday

In 1981, just three decades ago, editor HRM began a series of experiments to identify what 'messages' for toothpaste might persuade customers of Colgate toothpaste to remain with the brand, and which might entice prospective customers currently buying the competitor toothpastes to 'switch to Colgate. This rather simple question and the ensuing experiment, run in Canada at the behest of the late Court Shepard, General Manager of Colgate Canada, would turn out to be the basis of a science. It's hard to think of a science built on commercial products, dealing with the issues of everyday life. But this volume is testament to that science.

Our ingoing assumption at that time was that the consumer may not know what he wants, but he will know it when he sees it. We had another assumption, just as important, but we did not realize it then. That assumption was that by giving a person a compound test stimulus, we would increase the chances of identifying what elements were working. That is, it wasn't a case of learn by 'isolating,' but rather learn by creating naturalistic mixtures of ideas, testing these combinations with consumers, and then deconstructing the reactions to these combinations into the contributions of the components.

For the first 30 years afterwards, from 1981 to 2011, we hummed along, doing these experiments, working for clients, answering business problems ranging from toothpaste to prescription drugs, from services in a bank to public policy about emergencies. Those were, formative years, filled with learning, insight, and of course excitement. The test methods now developed, we applied the new approach to various problems, including dividing people by the pattern of their responses to test stimuli, and more profoundly by the pattern of their utility values, the 'driving powers' of the individual elements. In all that work, the emphasis was on applying the method as a consumer-research technique to business issues.

Enter a Grander Vision – A Science of Everyday Life

Perhaps it's one of the unwritten rules of science; work at a problem long enough, and soon the problem moves from something momentary to be solved to a source of wonderment about how the world works. We had this experience with the research approach of this book, RDE, Rule Developing Experimentation. Yes, we had begun with the trite problem of communicating what to say to make consumers buy more toothpaste. What could be more prosaic; almost a throwback to the 1950's. Yet along the way something happened. The problem of discovering compelling messages grew to something else, to a deeper measurement of what motivates people to respond.

And so the science of RDE began. It had started with the trite, the simple, the everyday. But then, almost subtly, and after two decades, a shift occurred in the nature of the way RDE made itself useful. Over time the focus moved from finding the specific answer, the one or two messages, to themes, to general patterns of ideas that excited consumers. The year 2001 was a seminal year; often seems to us that around that time we realized that projects using RDE were as focused on understanding people in their everyday lives as they were in solving a business problem. For instance, studies on what to communicate to the California citizen about energy costs, a practical problem posed by a utility, turned out to be a study of the mind of the citizen with respect to energy issues. Simple fascination with the results, with the patterns emerging from nature, blurred the boundary between problem and solution, or between person and structure of beliefs. Reading the data tables turned out to be a detective story, ferreting out how people thought. Only secondary was the importance of the actual problem, the reason for the study in the first place.

As this subtle shift continued, we realized that this new approach was creating a science of the ordinary, the everyday life, step-by-step, almost insensibly building a structure of how we go about evaluating alternatives. We were aware of similar ideas in the marketing literature, of so-called tradeoff-studies, where the goal was to understand what the consumer felt to be important. However, we were heading in a different direction, creating an archival science of the ordinary, a science whose data could be housed in tables and books, and pulled down from the shelf at any time to get a better idea of how people think and what people value.

What is the Essence of this New Science

As we rushed headlong into experiments, studying all sorts of topics, from foods to utilities, from public policy to technology and technical services, we found ourselves facing the same question, again and again, asked by those who hired us to solve the problems, by colleagues who were interested in new 'techniques,' and most intriguingly by students at different universities where we lectured. The question simply was '*what is this thing called RDE, rule developing experimentation*?'

We editors, who had been schooled in classic science, were not accustomed to questions as profound as '*what is this new science*?' We were more accustomed to providing solutions as professional consultants and consumer researchers. Yet, when we gave our lectures at the universities, we could see the gleam in the eye of the students, who, perhaps even more than we, intuited that this RDE 'thing' was something bigger. Students, by the way, are the most critical; being so young they are often not particularly diffident, having no trouble asking the hard questions, and expecting honest answers.

So what is the essence of this new science? What is Mind Genomics®? Quite simply, it is the study of how people react to the world of their everyday. The goal of Mind Genomics® is to create a database of the ordinary, to dissect specific experiences (*e.g.,* buying toothpaste) into components, identify the different acts and messages, and then determine which of the components, acts, messages, drive consumer response, and which are irrelevant. At the end of the day, the vision is to open a book for any experience, show the experience dissected, show what's important and what's irrelevant, and finally identify different mind-sets, people who look at the same experience in different ways.

Reactions to this New Science

When we began publishing the results of our studies in the archival scientific literature, we found a number of reactions, many quizzical, some downright negative, all enlightening. Most of the research that we had grown up with over the past decades dealt with the 'grand questions' about some aspect of how the world worked. The typical scientific paper would begin with a hypothesis, a speculation about the relation between variables, a speculation typically grounded in the knowledge and scientific contribution of previous researchers. The research should be organized in a way to prove or disprove the hypothesis. The statistics came in one general flavor, inferential statistics, statistics to confirm or disconfirm sameness or difference between two observations: Did the observed data differ from what was expected, and at what confidence level?

Mind Genomics® and its tool, RDE, came to the scientific world with a different world view. Rather than hypothesizing how the world 'might work,' Mind Genomics® offered the organizing principle that it was here to 'map the dimensions of experience.' There was no need to offer a hypothesis, and spend the experiment proving or disproving that hypothesis. It was sufficient to identify a topic area, *e.g.,* energy policy, map out the different ideas in this topic area through phrases, and then determine which of the different ideas appealed to consumer respondents *versus* which turned them off. The result was a description of a small corner of the consumer mind. Mind Genomics® mapped that corner, identifying how ideas worked in the consumer's head.

It took quite a while to realize that the reactions to Mind Genomics® by other scientists were those of academics/researchers reared in the world of hypothetico-deduction to the notions and ideas offered by inductive research. We ended up realizing that Mind Genomics® was essentially organized induction; we would map out a corner of daily experience with ideas, learn 'what worked, what did not,' and then make more sweeping statements about how people react to the specific topic area. Even our segmentation of mind-sets into different groups was purely empirical and inductive; we identified people who showed different patterns of what was important, and from that information we speculated about the distribution of different 'mind-sets' in the population. All in all, inductive science, mapping new worlds, seeing what was out there in this exciting world of the everyday.

Inside the Scientist's Mind(s)

Developing Mind Genomics® has taken 30 years, a half a lifetime, three decades of experiences. In light of the nature of this book, a compendium of applications from many colleagues who have graciously said yes, we thought that it would be a good idea to share one's feelings about this journey. We're not talking about the science or the technology, nor are we talking about the applications or the future. We leave that to our contributors. Rather, we're talking about the inside of the mind of the scientist, how it feels to be on this journey, what it means from a personal point of view, the 'soft stuff' of science that's often forgotten in the rush to publish. So here are some of the feelings from the 'inside,' where it's happening.

The first feeling is slight astonishment that one could actually be part of something like Mind Genomics®. Most of us who do science or technology, whether basic or applied, grow up with the idea that we are going to 'add' to a corpus of knowledge, that we are going to be part of a community of scientists, doing normative research, identifying promising areas, adding to the literature, but doing so in a less than dramatic manner. We may dream of winning prizes, the Nobel, for instance, but we realize that that's probably a pipe dream. Coming face to face with the implications of Mind Genomics®, is in the colloquial expression, 'something else, entirely.' We recognize that this new science, flawed but promising, given birth with so much excitement but also trepidation, may be something important. That feeling alone is what astonishes us, and continues to astonish as we make new discoveries.

The second feeling is curiosity. Just what world have we opened up? We don't know. Each new study is an adventure. The sheer simplicity, ease, speed of doing RDE studies and adding to the corpus of this new science makes us want to explore. Every topic we read about, from digital piracy to food safety, from stock market investments to creating new political parties, ends up being the inspiration for a new study. The goal was to find out – just how does that part of reality 'work?' And, that the curiosity, the inner energy of the scientist.

The third feeling is gratitude, an inner joy that comes from knowing that one has contributed something, perhaps of great value, to the generations to come. We all want to leave something of ourselves to the world, to those who we may never see, but who will carry on after us. Mind Genomics®, the science, and RDE, the tool, is our contribution to the next generations. We don't know where the science will take us; prophecy isn't one of our gifts. But, we do know that we are giving those who follow some new tools that will move science and the world forward. And what could be more delightful than having our colleagues, the authors of chapters in this book, join us in giving to the next generation. For all that we are grateful.

As we close this foreword, we'd like to take a moment to thank some of the people who helped us along the way. No work of science, especially one that takes 30+ years to develop and ripen, happens in isolation. There have been many, starting with the late Court Shepard who had faith in us, moving to the many computer programmers who worked with us on the algorithms, and on to the staff of researchers at Moskowitz Jacobs, Inc. in Westchester County, New York, who struggled with the science, applying it to client problems, and in so doing improving it ever so much, continually, and with good spirit.

Our thanks go to our colleagues who accepted our invitation to write chapters. We are grateful to them; they bring so many different, new perspectives that we feel them to be co-inventing this science with us. Each colleague's point of view sharpens the message, moves our thinking forward, and moves the science into new realms.

Finally, no work comes about without those who support it, day and day out. We'd like to thank our editorial coordinator, Linda Ettinger Lieberman, for her years of work, pulling the book together, making sure that all the details were taken care of, and essentially freeing us to do science while she made sure that we 'delivered the goods.' Thank you Linda, for this effort, and just as much for all the other efforts you have made on behalf of Mind Genomics® and RDE. Your efforts have truly helped move this science forward.

Alex Gofman, Ph.D. (Deceased)

&

Howard R. Moskowitz, Ph.D.

Maskowitz Jacobs Inc.
White Plains, New York
USA

List of Contributors

Ashman, Hollis	Muse LLC, Powell, Ohio USA
Beckley, Jacqueline	The Understanding and Insight Group, Inc., Denville, NJ, USA
Bevolo, Marco	Independent consultant/author, The Netherlands
Ewald, Jeffrey	The Optimization Group, Ann Arbor, MI, USA
Ettinger Lieberman, Linda	Moskowitz Jacobs Inc., White Plains, NY, USA
Fyrbjork, Johanna	IPOS Sweden, Stockholm, Sweden
Gabay, Gillie	College of Management Academic Studies, Tel Aviv, Israel
Gacula, Maximo (deceased)	Gacula Associates Consulting, Scottsdale, AZ, USA
Gadd, Jennifer	Gadd Research Inc., Toronto, Ontario, Canada
Gadd, Mike	Gadd Research Inc., Toronto, Ontario, Canada
Gassmann, Oliver	University of St. Gallen, St. Gallen, Switzerland
Gofman, Alex (deceased)	Moskowitz Jacobs Inc., White Plains, NY, USA
Huizingh, Eelko K.R.E.	University of Groningen, Groningen, The Netherlands
Jeavons, Andrew	Nebu, Cincinnati, OH, USA
Lopetcharat, Kannapon	Nouveau Centric Co. Ltd. Bangkok, Thailand (Formerly with Avon Products, Inc., Suffern, NY, USA)
Mets, Tõnis	University of Tartu, Estonia

Moskowitz, Howard R.	Moskowitz Jacobs Inc., White Plains, NY, USA
Narasimhalu, Arcot Desai	Singapore Management University, Singapore
Paredes, Dulce	Takasago International Corp., USA, (Formerly with Avon Products Inc., Suffern, NY, USA)
Rabino, Samuel	Northeastern University, Boston, MA, USA
Ray, Ipshita	Pace University, New York, USA
Shofu, Rieko	Hakuhodo, Consulting, Tokyo, Japan
Teerling, Marije L.	Telematica Instituut, The Netherlands
Zace, Sokol	Contact Design Inc., Boston, USA
Zeschky, Marko	University of St. Gallen, St. Gallen, Switzerland

PART I

THEORETICAL FOUNDATION OF RULE DEVELOPING EXPERIMENTATION

CHAPTER 1

Origins of RDE and the Role of Experimentation in Consumer-Driven Innovation

Alex Gofman*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: A key to business success lies in consumer-driven innovation. Rule developing experimentation (RDE) is a systematized solution-oriented business process of experimentation, which designs, tests and modifies alternative ideas, packages, products, or services in a disciplined way using statistical design. RDE uses either conceptual or physical prototypes. RDE applies to new product development, but can apply to more general social issues beyond the realm of products. RDE stems from the consumer-driven proactive approaches to structured experimentation, focusing on consumer preferences. RDE as implemented on the concept level uses so-called partial profile conjoint analysis. RDE's test stimuli often comprise incomplete concepts or vignettes created according to a specific type of experimental design (isomorphic permuted experimental designs). RDE uncovers pattern-based latent segments, as well as revealing the nature and magnitude of explicit and implicit interactions between the pairs of stimuli that RDE studies (so-called synergism and suppression). RDE traces its origins to experimental psychology, as enhanced through the driving power of business and social science. When applied properly, the developer and marketer discover rules and patterns defining what appeals to the customer, even in situations when the customer can't articulate the need, much less the solution.

Keywords: Consumer-driven innovation, consumer preferences, new product development (NPD), experimentation, rule developing experimentation (RDE), conjoint analysis, experimental design, regression analysis, fractional experimental designs, individual designs, dummy variable regression, incomplete concepts, interactions, pattern-based segmentation.

CONSUMER-DRIVEN EXPERIMENTATION

In past decades, consumers have become increasingly involved in the innovation process, a major factor of today's business success. The research of von Hippel (1986, 1988, 2005) and von Hippel and Katz (2002) points to the critical importance

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

of consumer involvement in innovation, particularly in product development. Sanchez and Mahoney (1996), Berghman, Matthyssens and Vandenbempt (2006), Gold, Malhotra and Segars (2001) and Möller (2006) demonstrate that using consumer and market knowledge can lead to an increased effectiveness of innovation process.

Innovation that is facilitated by organized and structured information, especially about the market and the consumer, is a vital resource to drive business success. So argue Inkpen and Dinur (1998) and Li and Calatone (1998), respectively. As Drucker (1995) pointed out, "…knowledge has become the key economic resource and the dominant—and perhaps even the only—source of competitive advantage… Such knowledge includes, but is not limited to, knowledge about one's company, industry, competitors, customers".

In his works, von Hippel (1976, 1978, 1986, 1988, 2005) demonstrates that a key to success lies in knowledge-driven innovation, fueled by understanding the consumer. Studies by Hurley and Hult (1998), Calantone, Cavuşgil and Zhao (2002), De Luca and Atuahene-Gima (2007) and Walter, Lechner and Kellermanns (2007) further affirm the importance of knowledge accumulation, reuse and transfer, respectively, as innovation drivers.

Analyzing consumer-driven knowledge for businesses, scholars such as Deshpandé and Zaltman (1982), Davenport and Klahr (1998) and De Luca and Atuahene-Gima (2007) conclude that such knowledge represents a firm's resource. The knowledge enjoys the potential to affect a firm's marketplace position. Ruggles (1998) and Schlegelmich and Penz (2002) confirm these findings based on the analysis of multiple commercial implementations.

In a recently released worldwide IBM study, the majority of the executives reported that they pay *extreme attention* (italics mine) to understanding consumer preferences in their product development. Such attention, in turn, points toward an even deeper involvement of consumers into the process. The effect—the *prosumer*—a consumer/producer deeply and extensively integrated into the value chain. The *prosumer*'s function? To enable and promote more precise customization of the production processes. Consumer-driven knowledge is critical there, concludes IBM (2008).

New product development (NPD) and related business processes are becoming ever more important in our evolving knowledge-based economy. Hunt and Morgan (1995) point out that scholars and increasingly companies acknowledge the significance of NPD as a factor in establishing competitive advantages vis-à-vis rivals. The attention to NPD is growing, gaining ever more significance. Leonard-Barton (1992), Olson, Walker and Ruekert (1995), Hunt (2000), McDonough, Kahn and Barczak (2001), Belliveau and Griffin (2002) and Trott (2008) progressively build and then reinforce the notion that fast, efficient, targeted and disciplined innovation in NPD is crucial for the economic success of the modern enterprise because it allows firms to occupy marketplace positions of competitive advantage. That, in turn, argue Ayers, Dahlstrom and Skinner (1997), results in the improved financial performance reported by many firms.

According to Riquelme (2001) and Thomke (1998, 2001, 2003), new product launches are a key driver of business growth. Jaeger, Rossiter, Wismer and Harker (2003) argues that in today's business environment, a continual supply of novel products is essential to retain one's competitive advantage. At the same time this continual supply doesn't always have a happy ending. The harsh reality, argues Flavin (2008), is that a majority of these new product launches simply fail. Across industries the failure rate for new product launches is dismal, ranging from an estimated 50% to 75% (ibid). The failure rate is higher than this dismal 75%. Food is a good example where failure is, by far, the norm. Goldman (2005) reports that regardless of a company's well-publicized intent to bring innovative food products to the market, the majority of the new launches are just replicas, fundamentally copies of what the competition is already featuring. This in turn results, for example, in the new food products failure rate of approximately 72–88% as reported in Buisson (1995), Rudolph (1995) and Lord (1999). It's not for lack of trying. Fuller (1994) estimates that for each food product that goes into the test market, 13 products are developed in laboratories and tested and rejected, before the one prototype that survives goes forward to market.

One of the ways to reach these business goals uses the strategy of *experimentation.* Experimentation, although often thought to be only in the domain of science, actually lies, according to Thomke (2001), at the very heart of NPD. Experimentation thus connects to corporate values, habits, strategies and organizational structures. In a later paper, Thomke (2003) further points out that experimentation, although an essential part of NPD and new product launches, has only recently come under scholarly scrutiny. Thomke (2001, 2003) also points to the importance of doing experiments fast, quoting Edison, who believed that the real measure of success is the number of experiments that can be conducted in 24 hours. Although the latter is an extreme opinion, it underscores the increased pressure of the competitive environment.

Businesses can improve and compete through experimentation. It's not just experiments alone, however. It's the tools to help experiments occur. These experiments often are intractable because they are complex, require a lot of time, or are simply not done because they intimidate. The experiments may not work because they are long and complicated (*e.g.,* adaptive experimentation (Sasena, Parkinson, Reed, Papalambros and Goovaerts, 2005)), or simply don't work because they don't get the right consumer inputs at the beginning or don't create actionable results at the end.

There are two competing approaches to experimentation: *unstructured* and *structured*. Despite massive theoretical work, Ziman (2000) and Trott (2008) write that frequently corporations use unstructured experimentation in their NPD efforts. Researchers from the Wharton School of the University of Pennsylvania and University of Texas, respectively, Wind and Mahajan (2001), analyze the use of unstructured experimentation at Seiko, which develops over 2500 watch designs every year and introduces them to test markets. The successful prototypes are further fine-tuned, retested and then launched in the target markets. The experimentation is not systematic but is extensive. Belson (2003) further studied the applications of this approach to mainly unstructured experimentation at Sony, which develops, tests and measures about 1500 product prototypes annually. About 20% of them are completely new designs. Only a small portion of those designs find their way to the global market. This massive scale of what turns out to be unstructured experimentation is feasible, according to Balasubramanian, Krishnan and Sawhney (2001), because prototyping is inexpensive and one can make many prototypes economically. Yet this type of experimentation isn't systematic and does not generate rules to create corporate knowledge. There is evolution, but not necessarily residual information of a formalized nature.

Despite the fact that systematic experimentation is more efficient than just undirected experimentation, it is taking a concerted effort in corporations to drive home this message. Academics recognize the value of such experimentation as being more efficient and thus getting to the goal more quickly and less expensively (Dahan and Mendelson, 1998, 2001). This recognition of the value brought by what we might call "systematics" resonates well with Thomke (1998, 2001, 2003) who demonstrates methods to lower the high cost of experimentation in corporations as a way to stimulate innovation. Thomke's works continue to argue the point and bring home the evidence that new technologies that facilitate experimentation make it easier than ever to learn from complex experiments, in a systematic fashion and of course do the learning within budget, *i.e.* inexpensively.

Expanding experimentation goes beyond knowledge building. The expanded use of experimentation may well provide an opportunity to take innovation to a new level. However, it is not just experimentation *per se*. It's also the way the experimentation is done, among whom it is done and how the information is integrated. Thomke (2001) stresses that corporations must be amenable to be willing to rethink R&D from the ground up. Such rethinking, according to Beckley, Foley, Topp, Huang and Prinyawiwatkul (2007) and Kantowitz, Roediger and Elmes (2008), involves the consumer, who must test the new propositions at the level of concept.

Yet it's not just the consumer and the testing. Rethinking involves a *systematic assessment* of consumer preferences at the level of concept. We are not talking here of tactical issues such as line extensions, which often don't need product trials. One can get by with concept work alone. For instance, Dickinson and Wilby (1997) suggest that in many cases, such as the tactically oriented project involving product line extension, effective concept tests do not even require product trials. Further developing the field, Lees and Wright (2004) find that a respondent's answers to attitude and purchase intent questions show only minor variation with different formulations of the concept test statement and that the ranking of the concepts shows no substantial changes across the different formulations.

Rather, we are talking here of concept testing in the much larger sense of experimentation. We are talking about systematic experimentation using concept testing to understand the "algebra of the customer mind". This direction is supported by Moskowitz, Porretta and Silcher (2005) and Sinkula, Baker and Noordewier (1997), who advocate innovation at the concept level (as opposed to physical prototyping). The innovation at the concept level uses experimentation. The innovation at the concept level is knowledge-based and underlies a subsequently disciplined development strategy, which produces a database that simultaneously helps to understand and to create.

Experimentation is also important because of the nature of people, our consumers. Millett (2006) and Klink and Athaide (2006) criticize some methods of obtaining consumer preferences. The evidence is clear. All too often consumers cannot articulate exactly what they need, want, or like, even when asked directly. They may repeat advertising buzz-words but can't design what they want. Agreeing with such arguments, Green and Wind (1973) and Hauser and Rao (2003) point out that it is very difficult for consumers to articulate their needs and desires and researchers have to use other means to understand their motivations. Some arguments, such as presented in Kiley (2005), go so far as to conclude that focus groups really may not provide a reliable direction for NPD or message optimization. Krieger, Green and Wind (2004) point to a solution for this problem; experimentation by presenting consumers with a set of systematically designed concepts. Research shown in Moskowitz *et al.* (2005) proves that it is much easier for consumers to choose a preferred option from a set of already executed, immediately available, product and/or positioning concepts. The experiment makes all the difference. It forces one to create the alternative. Innovation is merely selecting from among that which could be and now actually exists.

CONJOINT ANALYSIS AND CONSUMER SEGMENTATION

Krieger *et al.* (2004) wrote that that conjoint analysis has become increasingly prevalent as a major approach to studying consumer preferences. Even without their well-documented statement one need only search the terms conjoint analysis and conjoint measurement using Google®. The popularity of conjoint measurement is partly attributable to its basis in experimentation. Arorar and Huber (2001) argue that the experimentation afforded by conjoint analysis enables researchers to model choices in an explicit competitive context. By so doing the researchers believe that they realistically emulate market decisions. In addition, Atkinson and Donev (1992) point out that conjoint analysis is becoming a major tool in concept development, increasingly used for industrial-oriented NPD (Orme, 2006). The history of conjoint analysis is not business but rather the axiomatic measurement theory of the type beloved by mathematicians and mathematical psychologists. In its original form, conjoint analysis was first introduced as an intimidating but somewhat entrancing approach to measurement by Luce and Tukey (1964) and published, mathematical notation and all, in the first issue of the *Journal of Mathematical Psychology*. The reality was that conjoint measurement extends the idea of functional measurement summarized in Anderson (1977). It was Anderson who averred in far simpler, more inviting prose that the researcher could learn a lot by studying combinations of stimuli and applying them to the mixture a decomposition rule, such as analysis of variance or regression analysis.

In conjoint analysis, products are defined in terms of possessing a limited number of relevant attributes or characteristics. These relevant attributes comprise at the highest level categories or silos of features or emotional attributes of new products, with each category in turn containing a limited number of levels (elements). The "products", often called profiles, concepts, or vignettes (all are the same) have a known composition. In turn, respondents evaluate these concepts, using some type of scale. In some versions of conjoint analysis, the respondents rank the vignettes. In others the respondent chooses one vignette from a group.

As Louviere (1988) explains, conjoint analysis comprises a decompositional approach that analyzes consumer preferences. Respondents provide overall scores to a concept (product profile). The conjoint analysis imputes the individual preference contribution for each component of the concept. The overall utility of a product profile can be reconstructed by adding together the separate attribute preferences values (impacts or utilities). Going one step further, this type of arithmetic generates what we call the *compensatory preference model*: "low" scores on a certain attribute can be compensated by a "high" score on another attribute. There are, of course, other "flavors" of this conjoint approach. For instance, Vriens (1995) presented the notion of *noncompensatory preference models.* This latter version of conjoint analysis posits that certain attributes must have a minimum or maximum level before a profile is considered attractive could be utilized as well.

Whichever approach one subscribes to, compensatory or noncompensatory model, the marketing world has warmly adopted conjoint analysis. The approach makes sense. And it delivers the necessary information. In the marketing world, conjoint analysis is best known for being a research technique by which one can investigate combinations of features to identify which combination is best. Developed by Paul Green and his colleagues at Wharton School of the University of Pennsylvania (Green and Srinivasan, 1981), conjoint analysis is now widely used in high profile projects to design products and services in a variety of different categories. Reviewing the history of conjoint analysis, Krieger *et al.* (2004) conclude that since the early 1980s, conjoint analysis has evolved into industry's most widely applied marketing research tool to measure the multiattribute utility functions residing in the "mind" of the buyer.

We can discover the intellectual and mathematical foundation of conjoint analysis in the very large technical domain known as *experimental design.* Experimental design in its most simple form comprises a statistical plan that lays out the specific combinations of the concept elements (Atkinson and Bailey, 2001). The traditional experimental design creates one single set of combinations, which is evaluated by the different respondents in the study. The analysis may be at the level of the individual or the level of the group. The fact that there is only one basic design means that the different respondents, *i.e.* the increasing number of cases, serve as a way to develop a stable mean for each one of the combinations created by the design. Other points of view hold that the single experimental design can be used as a kernel. The kernel is then permuted to create different, so-called isomorphs of this basic kernel. Each respondent evaluates test stimuli from one of the many isomorphs created for this purpose. Finally, although many researchers like to have the respondents evaluate combinations (test concepts) comprising exactly one element from each category, there is good statistical reason to have respondents evaluate incomplete test concepts, occasionally lacking one or more categories (Moskowitz *et al.* 2005).

Conjoint analysis does not proceed automatically, however, despite its elegance and power. There must be the criteria of independence, readability and in the end "actionability". In the process of evaluating systematically created combinations, the elements of the concept constitute the independent variables and the attribute ratings become the dependent variable. Hair *et al.* (1995) point that the concept elements must be statistically independent of each other or else the criteria to create an additive model by regression would be violated. At the very practical level the concept must be readable. To be useful for NPD, these elements have to be meaningful within a concept and ideally realizable in an actual product.

We now turn to the nature of the test stimuli. These determine the nature of the results that we get, their interpretation and ultimately their generalizability across studies and across time. There are two classes of stimuli: full profile and partial profiles. They lead to rather surprisingly different results.

Full-profile conjoint analysis presents the test concepts that always contain exactly one element from each category. Krieger *et al.* (2004) conclude that this format for the test stimuli is the relevant one when analyzing the nature of choice for purchases when a person buys a single product. Full-profile conjoint analysis shows full product designs, which replicates the way products are actually purchased. Full-profile conjoint can generate individual models, at the level of a model for each participating respondent. This resultant model enables the marketer to identify the (sometimes a) "most preferred" product for each respondent. Despite the real-world nature of the full-profile approach, Green and Srinivasan (1981) point to a key drawback–the study must be limited to approximately six attributes with each attribute comprising no more than five to nine levels. This limitation does not allow the research to describe highly complex products. Furthermore, existing full profile approaches utilize the same underlying experimental design for all respondents. Any one design thus studied many times across respondents by necessity comprises a relatively small number of variations of the test concepts. One unexpected and potentially disturbing consequence is that it is virtually impossible to detect interactions between pairs of elements. There are simply too few combinations of elements for these interactions to be revealed through standard statistical tests. In addition, Krieger *et al.* (2004) demonstrate that full-profile conjoint analysis, with one element from each variable always present, generates collinear models. The utility values end up being relative to each other, rather than absolute, making it impossible to compare utilities across the different attributes. The utilities from full-profile conjoint analysis can only be compared *within* an attribute, significantly lessening the attractiveness of this particular version of conjoint analysis.

Partial-profile conjoint analysis adopts a different point of view (Krieger *et al.* 2004). Some features are explicit, presented in the test stimuli, whereas other features are assumed to be constant throughout the entire interview. The number of the elements in the test stimulus may vary from two to many. Kuhfeld, Tobias and Garrat (1994) argue that partial-profile methods are ideal for computer interviewing, explaining some of the popularity. Today's approaches to partial-profile conjoint analysis typically use a single (sometimes randomized) experimental design. The effect of such a design is to make the study setup easier for the researcher. The statistical outcome is to reduce the variations of the concepts and elements combination. The unforeseen outcome is that today's typical designs make it difficult, often impossible to detect true interactions between pairs of elements (*i.e.* synergisms and suppressions), as there are not enough stimuli in the proper format to estimate linear and interactive effects.

Discrete choice conjoint analysis represents yet another variation of today's conjoint analysis. In discrete choice, the respondent is presented with a set of different "profiles" or test concepts, presented simultaneously on a computer screen or on a card. The respondent's task is to choose the concept most acceptable. Discrete choice studies trace their intellectual heritage to paired comparison methods in psychology, where the ingoing belief is that the respondent is best at selecting among alternatives rather than acting as a measuring instrument. Discrete choice methods are most frequently utilized for products where consumers purchase multiple products distributed over many brands over the course of a year, doing so in proportion to the relative desirability of those products. According to Green and Srinivasan (1978) and Krieger *et al.* (2004), the advantages of discrete choice methods includes their tolerance toward perceptional incoherency of the concepts; flexibility to direct combinations of the elements with selected constraints (*e.g.,* showing certain features with one brand and a different set of features only with another brand); and some ability to measure interactions between two attributes, such as between price and brand. At the same time, discrete choice is less applicable to NPD than other methods. Moskowitz *et al.* (2005) demonstrate that discrete choice data can only be analyzed at the aggregate level. The unforeseen consequence is the inability to segment consumers in a tractable way based on the patterns of their individual utilities. Furthermore, discrete choice methods do not readily generate accurate estimates of market share because the data cannot be adjusted easily to reflect the different awareness and purchasing patterns of the individual respondents participating in the study.

Beyond the statistical properties of the conjoint methods lie the issues of scope and application. For many problems, it suffices to work with few stimuli that can be deeply investigated in a conjoint task. On the other hand, as researchers and marketers become increasingly facile with the procedures, there is the inevitable desire to increase the scope, to test more stimuli in a study and to expand the scope of what is tested (from words to pictures, for example). According to Green and Srinivasan (1990), early applications of conjoint analysis tested relatively few concept elements. These early studies focused on product design and so the elements in the study generally were the more rational features found in products.

Over time and with the increasing acceptance of conjoint analysis as a test method, there arose the issue of increasing the number of elements to be tested in the study (*e.g.,* from a dozen to a few dozen and then to 100 to 300). There also arose the issue of creating a model for many elements at the individual respondent level.

One traditional way to deal with many elements instructs each respondent to select the specific concept elements that would be most appropriate to him and then to discard the rest so that they are not tested in the conjoint portion of the experiment. This procedure requires that the respondent explicate importance before participating. Having a respondent first evaluate the individual elements prior to testing them in combination is known generally as *hybrid conjoint analysis*. Selecting only those that are deemed important when evaluated alone is known as *adaptive conjoint analysis* (ACA) (Johnson, 1984). Toubia, Simester, Hauser and Dahan (2003) suggest that ACA works best when all of the elements to be considered are similar in terms of quality. By this we mean that the elements are similar, such as simple statements about product functionality that can be judged alone and accepted or discarded as being relevant or irrelevant.

Although one might think that adaptive methods solve the problem by focusing only on relevant elements, there are drawbacks to such direct methods. Krieger *et* *al.* (2004) considered three problems with such adaptive approaches that make ACA a less than desirable approach to deal with a large number of concept elements. These are effort (substantial up-front efforts for respondents), unnatural decision process (forcing respondents to make difficult rational decisions on individual element level rather than a combination of the features) and subsequent analytical weakness (inability to create individual models).

Until recently, most applications of conjoint analysis had to work around various limitations that weakened the method. They either presented the same set of concepts to all respondents, or blocked them across sets of respondents. The concepts could be chosen randomly, or chosen based on responses from prior respondents. Yet the importance of individual models has always been recognized as the "gold standard". Adaptive design for individual respondents was first considered by Toubia *et al.* (2003) and Toubia (2004) in metric paired-comparison settings. This new approach, named the polyhedral method, works by iteratively constraining the polyhedron of feasible subutility (part-worth) values. As the problem is computationally hard, many approximations, such as *Q-Eval* proposed in Iyengar, Lee and Campbell (2001), are employed. Toubia *et al.* (2003) and Toubia (2004) extend this technique to metric paired-comparison queries in. At the same time, the polyhedral method solves only one of the three intrinsic problems for ACA, the lack of individual models.

Beyond its capabilities as a method to deconstruct ideas and messages into the contribution of the separate components, conjoint analysis provides a way to deconstruct the environment into components and by so doing offer a way to detect "weak signals". A weak signal is a small movement in a variable that might well become far stronger and turn into a trend. The exercise of setting up a conjoint analysis experiment makes the discovery of such weak signals more likely, because of the following reasons:

- By forcing the researcher or ideation expert to edit the elements, better thinking emerges. Homework, working out the language, forces clarity and gets rid of a lot of the fuzziness. According to Green and Srinivasan (1990), the concept elements have to be simple, stand-alone ideas, phrased in the active case. When just emerging from ideation, frequently the idea is poorly expressed. Flores, Moskowitz and Maier (2003) suggest that the discovery and polishing exercise itself is as valuable as the ultimate discovery of weak signals. Moskowitz *et al.* (2005) point out that during the course of preparation from ideation to evaluation in the conjoint study, the elements themselves mature from simple notions to better-expressed statements.

- Darwinian principles operate. The performance of ideas is judged against different backgrounds, encouraging "survival of the fittest". Conjoint analysis uses experimentally constructed concepts comprising several different ideas. Kantowitz, Roediger and Elmes (2008) argue that to the degree that a concept element performs well in multiple backgrounds, one can be sure that the idea is good. According to Wind and Mahajan (2001), this rigorous test speeds up the development process because it creates an objective method to identify good ideas.

- Conjoint analysis leads to segmentation, which may reveal the great promise of an idea, otherwise masked among the averaged data. Segmentation is key to achieving increased acceptance because consumers do not share the same preferences. As pointed by Green and Krieger (1991) and Beckley *et al.* (2007), what appeals to one consumer may not appeal to another.

The statistical infrastructure of conjoint analysis generally devolves to regression analysis. Regression analysis reveals the quantitative nature of relations among variables. When properly executed, regression analysis ends up with an "ah-ha" moment, as relations emerge that lead to insight. Scholars (*e.g.,* Bretscher (1995)) point out that the computational machine underlying conjoint analysis is often the familiar *ordinary least squares* (OLS) regression. OLS offers a simple, yet robust method of deriving alternative forms of respondent utilities (part-worth, vector, or ideal point models).

According to Bretscher (1995), OLS confers a number of advantages on the analysis of conjoint data. These are simplicity, the intuitive meaning of the results and the capability of creating an individual level model by running the OLS on the data from a single individual. Reviewing many other methods, Harrell (2008) concludes that the ability to implement designs having larger numbers of attributes and levels (through fractional factorial designs described below) has made OLS the *de facto* standard for conjoint analysis.

Rao *et al.* (1999) formulate the objective of OLS in conjoint analysis, producing a set of additive part-worth utilities that represent the degree of acceptance/preference of each respondent for each level of a set of product attributes. In application, the OLS model uses a dummy matrix of independent variables, so each element is coded either 0 or 1. Each independent variable indicates the presence (1) or absence (0) of a particular attribute level. The dependent variable is the respondent's evaluation of one of the profiles described by the independent variables. This model is expressed as:

$$U(x) = const + \sum_{i=1}^{n} b_i x_i$$

where b_i are called *partial regression coefficient*s (parts-worth or impacts) of the independent variables *x*. They measure the variation in the value of *U(x)* that is due to a variation by one unit of the independent variables. Each individual variable generates its own partial regression coefficient. In the context of a conjoint analysis project with a purchase intent anchored question, the coefficients could be interpreted as a conditional probability of consumers being interested in buying this product.

We now move to the use of the method and the steps in the process. A conceptual map showing the different aspects of today's conjoint-based approaches appears in Fig. **1**. Fig. **1** illustrates the different aspects of conjoint as they are used in product development.

The knowledge development process using conjoint analysis typically begins by gathering initial ideas (features of the product, related emotional messages and pictures of components). The ideas could come from inside of a corporation as brainstorming, CRM data, or ideation. Today's technology and approaches provide a host of different approaches to gathering these ideas. In addition to traditional focus groups, the Internet is a new source of product ideas. The Internet itself provides different sources, such as blogs. There are other venues afforded by Web 2.0 including word-of-mouth (WOM) campaigns that serve as possible sources of ideas.

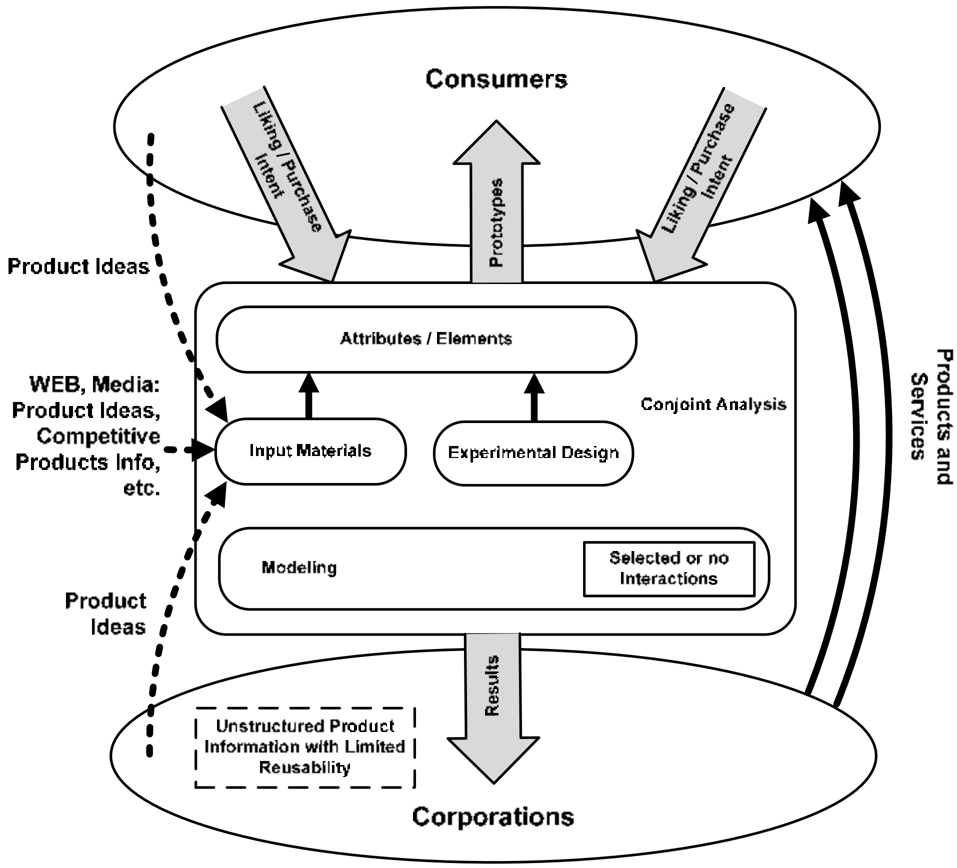


Figure 1: Conceptual map of the consumer-driven approaches to product development using conjoint analysis (source: Gofman, 2009).

The initial ideas are edited, often needing simple reformatting, so that they become concise, descriptive short sentences. The ideas are then classified into groups based on simple principles, such as the nature of the element (*e.g.,* type of product feature, how the product is used, *etc.*). For example, attributes and elements for a new cookie products could be (the elements are bracketed and follow the attributes): “Flavors” (“Chocolate-flavored”, “Orange-flavored”, "Strawberry-flavored"); "Size" ("Bite-sized", "Oversized", "Normal-sized"). The elements could be more descriptive and be presented in a textual or graphical way.

When creating the test elements one must always keep in mind that the conjoint methods are statistical and cannot do one's thinking. The old adage "garbage in, garbage out" applies here. It is important, therefore, to collect the elements, arrange the variables, create the test combinations and run the study in a way that guards against some well-known problems. Here are a few of the possible issues that one must watch for. None of these issues is particularly difficult to avoid:

- **Collinearity and thus no true estimate of the utility score of the individual elements.** Complete concepts do not allow estimating the absolute utility scores of concept elements due to multicollinearity (Krieger *et al.* 2004). The statistical requirement that the individual utilities sum equal to zero means that when a new element is introduced into the study, the utilities of the other elements must be readjusted because they have relative value. This readjustment means that one cannot use the results for reusing and databasing (absolute values of the utilities that could not be easily, if at all, compared across the attributes or projects (Beckley *et al.* 2007)).

- **Skewed results because of too few combinations**. Limiting the number of concepts, *e.g.,* in response to a desire to do less set-up work up front, could skew the outcome. In this limited set of combinations, the elements appear in just a few of the combinations. A few potentially very strong elements might distort the results (Rabino *et al.* 2007).

- **Insufficient variation of combinations** also prevents detection and estimation of interactions (pairwise and higher order Moskowitz *et al.* (2005)). Besides the individual contribution of the elements, some specific pairwise combinations might have an additional impact on consumer liking or purchase intent. Green and Devita (1975) point out the importance of researching such situations and estimating their utility scores. Chrzan and Orme (2000) point that simple methods for conjoint analysis ignore interactions on the premise that they are contained within elements and yet estimable. Over the years, increasingly more complicated methods have been introduced that partially address the issue of testing a limited set of interactions. Aiken and West (1991) suggest that a straightforward approach for estimation of all possible interactions—two-factor, three-factor and so on—would call for a full factorial design with at least two replications per cell. For realistically sized problems relevant to NPD, the number of combinations generated would be prohibitively large for respondent evaluation. According to Evgenoiu, Boussios and Zacharia (2002), estimating of all interactions using the existing approaches is so computation-intensive that it might need utilization of special high-performance support-vector computers usually found in the artificial intelligence area. They raise the need for an easy and reliable method to estimate how pairs of product features interact to drive consumer interest.

SEGMENTATION

Segmentation refers to the division of respondents by one or another set of criteria such that respondents in one segment are more similar on a set of criterion attributes. Neal (2003) defines market segmentation as the selection of groups of people who will be differentially receptive to a product. Segmentation of consumers into groups with similar preferences may provide an additional opportunity for product developers to create new and innovative products. Demographics generally do not account for the segments, nor do they predict the segment membership of any individual particularly well.

Well-defined segments are easier to target with specific products and market strategies. Gathering more relevant and predictive lifestyle and demographic information on the potential consumer allows the company to develop greater depth of knowledge about the targeted market and consumer, although as stated above, such information does not predict segment membership.

Moskowitz *et al.* (2005) point that in general, segmentation may be accomplished in two different ways—*a priori* segmentation and *a posteriori/post hoc* or latent-based segmentation:

In *a priori* segmentation, the segments are developed by certain theory-derived assumptions about the preferences of the respondent. Neal (2003) points that a large share of segmentation is done using exogenous variables such as geo-demographic characteristics, where the assumption is that people from different geo-demographic "breaks" may exhibit different utilities.

***A posteriori/post hoc* or latent-class segmentation** emerges from patterns of responses to test stimuli or responses in experiments. Examples of such stimuli used for latent-class segmentation are the individual utilities in conjoint analysis or the responses to a series of attitudinal questions, respectively. The utility values from conjoint analysis can and have been used as bases for segmentation. With conjoint analysis, one segmentation algorithm uses the pattern of utilities themselves, developing a distance measure between pairs of respondents based upon a distance metric such as *1-R,* where *R* is the Pearson correlation coefficient between two sets of utilities. The segmentation generates an interesting and often profitable way to divide a group of respondents. The patterns used for latent-class segmentation may be the responses to actual test stimuli, rather than the utilities from a conjoint study or patterns of responses for a questionnaire. For example, Pangborn (1970) shows a segmentation method using the pattern of liking of a test stimulus *versus* its sensory intensity. Pangborn's segmentation approach was first introduced in food research and only later adapted for concept studies.

RULE DEVELOPING EXPERIMENTATION

Rule developing experimentation (RDE) is a system for structured, consumer-based experimentation with conceptual prototypes, applicable to NPD and other areas such as public policy. To put RDE into perspective, RDE grew out of consumer-driven proactive approaches to structured experimentation. RDE focuses on uncovering consumer preferences at the concept level. RDE uses partial profile conjoint analysis, creating and testing incomplete concepts. These concepts are constructed according to a special type of fractional, main effects experimental designs (isomorphic permuted experimental designs). RDE uncovers pattern-based latent segmentation, as well as detecting and quantifying implicit interactions. Fig. **2** shows general categorization of RDE.

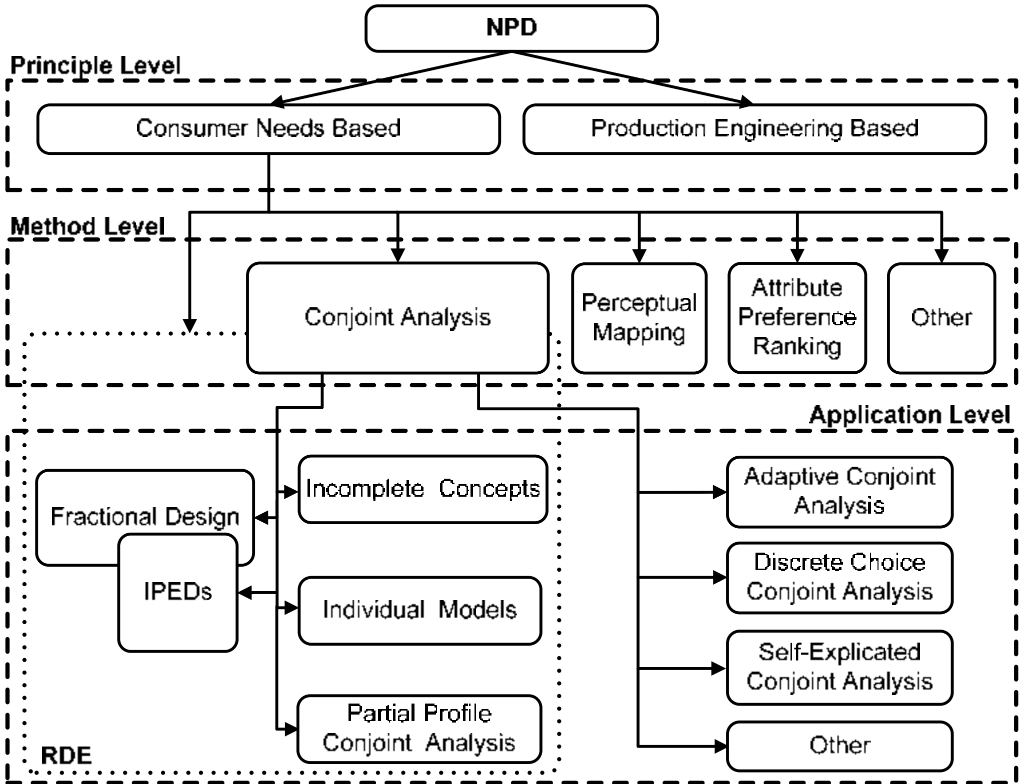


Figure 2: General categorization of rule developing experimentation (RDE) as applicable to new product development (NPD). RDE elicits and analyzes consumer preferences of the product prototypes at the level of the concept. Experimentally designed concepts are tested using the method of partial profile conjoint analysis. The approach employs a special type of fractional main effects experimental designs (isomorphic permuted experimental designs). RDE incorporates a number of features not currently incorporated into conjoint analysis (source: Gofman, 2009).

In its position in the bigger world of business, RDE represents a systematized, solution-oriented business process of experimentation that designs, tests and modifies alternative ideas, packages, products, or services in a disciplined way using experimental design, so that the developer and marketer can discover rules and patterns showing what appeals to the customer, even if the customer can't articulate the need, much less the solution (Moskowitz and Gofman, 2007).

The roots of RDE come from these three sources:

Experimental psychology. In essence, RDE is based on the direct Stimulus–Response model. RDE is founded on the realization that perception and behavior are linked in a two-way exchange. For instance, when you increase the level of sweetener in Pepsi Cola, it tastes sweeter. Liking can change as well; consumers can grow to prefer the sweeter cola. In fact, when you want to create the optimum Pepsi, one development strategy changes sweetener level, measures sweetness, measures liking and determines the sweetener level at which liking reaches the highest or optimum level. This example is a simple description of RDE applied to the problem of developing a physical product. By means of systematically changing the stimulus and measuring the response to that change, the investigator uncovers find the patterns or the rules.

The driving power of business. In the simplest of terms, businesses make products, offer services and generate profit. With increasing competition, businesses must always focus on opportunities to maintain their market share to survive. Thus businesses perennially look for opportunities to offer "new" (at least, perceived to be a fresh idea), "better" (according to the people buying it) and "profitable". Many decisions about what businesses offer are based on the predilections of so-called golden tongues, maverick executives who are hopefully truly talented individuals. For the other 99% of people, it is far more productive to learn how the world works and to discover the particular rules by which to make the offering better, make it cheaper and of course do it faster. Unless you are in that 1% of incredibly gifted or lucky predictors, business works better with rules. These rules reveal how to create winning formulations that taste great, messaging that "grabs" customers and packages, or magazines that attract customer attention and get purchased as a result. RDE is about how best to perform each of these tasks. RDE produces these results each time because of the systematic approach. The process takes just days, not years. Such speed and accuracy are vital for business.

Social science. Formal, scientific experimentation in the social sciences with the express objective of generating rules is just beginning. Not much has been done yet in the way that experimental psychologists and

> businesspeople do experiments. However, RDE is related to a field called adaptive experimentation (AE) or adaptive management. AE tries to find answers to ecological or social problems through trial and error, using feedback to drive the next steps. At each step in this process, the researcher looks at the data, tries to discern the pattern that might exist and then adjusts the conditions. The most publicized cases of AE are very lengthy, large-scale, even monumental, projects in the areas of ecology, theoretical science, or sociology/environment. AE does not, however, generate rules. Instead, AE searches for workable solutions using the process of experimentation. AE is not defined by a simple experimental structure with finite steps, nor is it governed by limited periods. In contrast, RDE follows a limited number of steps, in a limited time frame and then uses experimental methods in order to understand the algebra of citizens' minds.

Fig. **3** shows the conceptual map of RDE. The input materials are organized into attributes and elements, arranged into experimentally designed concepts. The process starts with the same sources of initial ideas as do traditional approaches. In addition, RDE often uses a data store containing structured information created by previous utilization of RDE, as RDE can accumulate, compare and update data across the projects. The initial ideas are formatted into concise descriptive snippets of text or other media. These become individual elements of the products, which for bookkeeping purposes are grouped into attributes of similar features.

The attributes (groups of elements) separate the elements into similar groups such as product features, benefits, or emotional messages. These elements are combined into the sets of concepts. The combinations are created according to a special type of experimental designs, *isomorphic permuted experimental designs* (IPEDs), which will be discussed in the next chapter. IPEDs generate individual (for each respondent) unique sequences of concepts, subject to statistically possible limits depending on the number of the categories and elements. With a sufficient number of respondents, an individual IPED allows the evaluation of every possible combination of the elements numerous times by different respondents within the framework of conjoint analysis. Such an approach will reveal higher-order interactions.

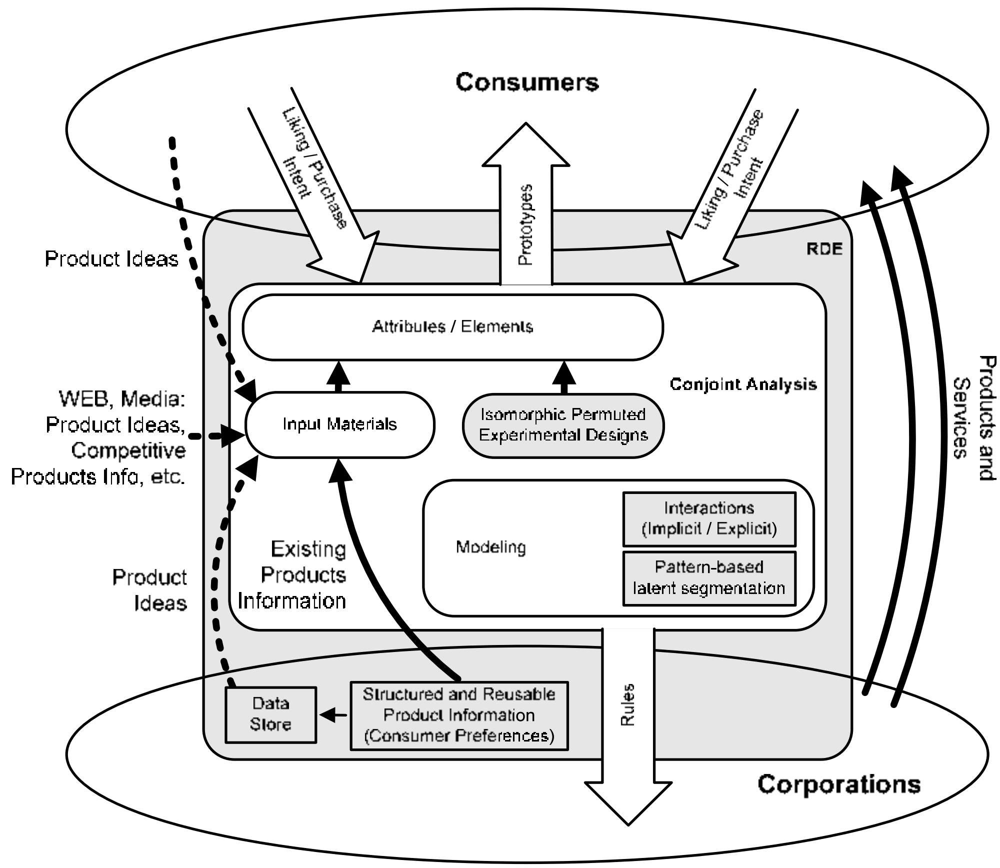


Figure 3: Conceptual map of rule developing experimentation (RDE) together with conjoint analysis. The proposed modifications to the conjoint analysis based consumer-driven approach to product development are highlighted (shaded areas) (source: Gofman, 2009).

The IPED creates the necessary combinations by which to identify and estimate pairwise interactions, as well as latent-class segments. Beyond that capability, the strategy of using IPEDs allows the regression modeling to estimate the *absolute values of the utilities* rather than being limited to estimate relative values. The latter constraint, with weaker estimation of parameters, characterizes the majority of today's methods.

The ratings assigned by a respondent are combined with that individual's IPED. The combination provides the necessary information to estimate the part-worth contribution of each element, by means of OLS. The outcome is sets of utilities, one set for each respondent.

The power of the individual utilities extends beyond the information which that information enables. As RDE now has the absolute values of the utilities for each element for each respondent, it becomes a straightforward statistical task to divide respondents on the basis of the patterns of their utilities. Thus RDE detects pattern-based latent segments based on combining individuals whose utilities are similar patterns. Green and Krieger (1991), Neal (2003) and Moskowitz *et al.* (2005) argue that such segmentation, based as it is on the pattern of responses to granular level, relevant stimuli, is more targeted and thus effective than are the traditional segmentation methods. After the segmentation has been done and the different segments identified as to what is most important, RDE then quantifies the most likely combinations of ideas (features or messages) that may be acceptable to consumers in each segment.

RDE STEPS

RDE encompasses six basic steps, which are similar to the steps usually used in conjoint analysis projects although in a modified form. The steps extend conjoint analysis to different fields:

Step 1. Prepare raw materials. Identify groups or classes of features that constitute the target product (offering, product, service). For example, in the case of a credit card offer, the tested attribute could comprise annual percentage rates (APRs) and rewards options. Every such attribute of ideas comprises several elements (different levels of APRs or various reward options).

Step 2. Create test concepts. This step mixes and matches the elements according to an experimental design to create a set of concepts. The second step is based on a variation of conjoint analysis (partial profile with incomplete concepts, based on individual models designed by IPEDs).

Step 3. Collect data from the consumers. This step constitutes the actual experiment with consumers, who are invited to participate and who evaluate the experimentally designed concepts. An example of a rating question: “How likely would you be to buy this product?”

Step 4. Analyze results. One key differentiating features of RDE is the individual model of utilities generated for each respondent. The individual models allow patterns to be discovered in the data, based on the individual elements and respondents. The models allow targeted optimization. The analysis uncovers all meaningful two-way interactions between pairs of elements. In addition, as RDE estimates the absolute values of the utilities, one can database the results to compare utilities within a study across variables and across studies.

Step 5. Identify pattern-based latent segments. As proved in Green and Krieger (1991) and Neal (2003), people's preferences transcend demographic groups. Using the IPEDs, RDE uncovers naturally occurring pattern-based latent segments of the population that show similar patterns of utilities. The approach of dividing people differs from the conventional ways that use gender, income, products purchased and the like.

Step 6. Apply the generated rules to create new products, services, offerings. RDE results can be combined into new products, which optimize some objective function, such as the sum of the separate utilities, including interactions, when those are known.

Fig. **4** shows the generalized model of the RDE process. The process begins by collecting the input materials for NPD. These materials may come from a number of sources such as corporations, consumers, databases of the existing features, competitive analyses, consultants with category knowledge and so forth.

First, the input material is sorted into groups (attributes) of related features. Afterward, an experimental design is applied to the elements. The design guides the mixing of the elements, creating individual sets of combinations. These individual sets of combinations are themselves actual small designs. Each individual design is isomorphic to the basic design, created through a permutation of the elements. Afterward, a computer-aided personal interviewing system or computer-aided Web interviewing system shows the different combinations of one person's design to that person and acquires the rating, on a combination-by-combination basis. Each respondent evaluates his set of combinations on one or several rating scales. On a person-by-person basis, the ratings are related to the elements using OLS regression. This regression analysis is feasible at the level of the individual respondent because the original combinations, one set per respondent, were designed to constitute a full experimental design. Finally, after the individual models have been estimated by regression analysis, the system identifies pattern-based segments through clustering and also identifies pairwise interactions (synergisms, suppressions).

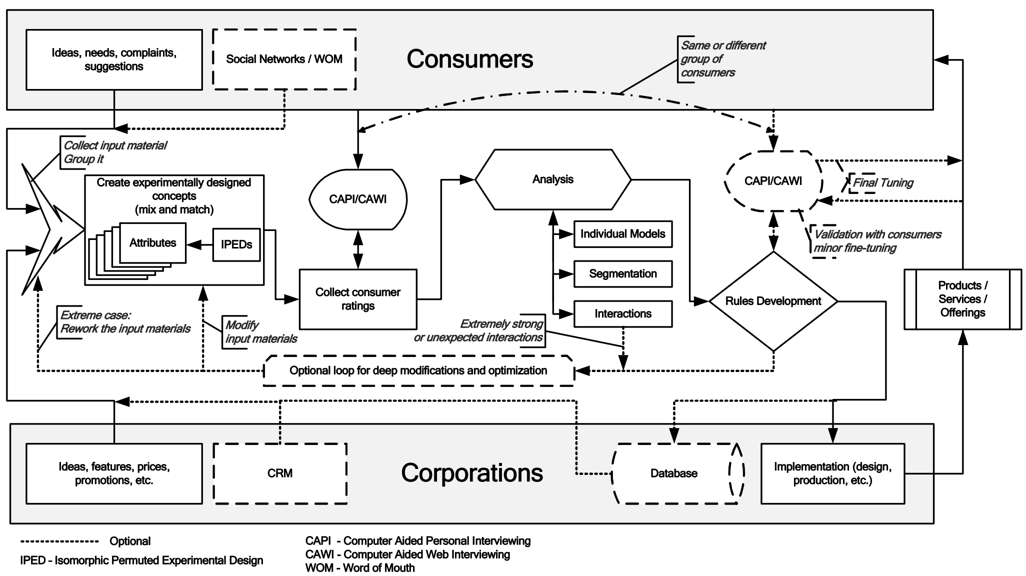


Figure 4: Rule developing experimentation (RDE) applied to new product development (NPD) (source: Gofman, 2009).

The RDE process does not necessarily stop after one iteration. The system may require a rework on the input material based on empirical findings during the course of a study. For example, some interactions between variables may emerge, shedding new light on the nature of the elements and suggesting other elements should be selected, especially when the interaction is suppressive so the combination performs especially poorly. In such cases, it is desirable to select new elements that don't exhibit this unexpected and undesirable way of interacting. The intermediate results could be optionally validated/fine-tuned with consumers. During this process, several optimized concepts would be tested with a new group of respondents and with modified elements, in order to fine-tune the results. The nature of such fine-tuning is left to the particular investigator, who might wish to invite the original group of respondents to evaluate the new test stimuli, or who might wish to work with an entirely new group of respondents. Optionally, the concepts could be analyzed in a qualitative session to better understand the consumer preferences. The goal of the iterations is, of course, the best combination of elements for NPD. The next step is implementing the rules for NPD. As iteration is always a possible necessity, in those cases when the results do not suffice to solve the problem, the investigator returns to the drawing board, sifts through the elements, discards the elements that do not work, tries other elements along with those that did work and proceeds to the next iteration. Experience suggests that by the time two to three iterations have been completed, the NPD effort will have identified mostly strong performing elements.

DISCUSSION AND CONCLUSIONS

Conjoint analysis, a class of research approaches based on experiment design, is used frequently in order to reveal consumer preferences in product development. One of the most significant drawbacks is the nature of the utilities, which have relative value, not absolute value. The problem comes from the specific design, which produces statistical multicollinearity. As a result, the utilities cannot be compared across variables in the same study and cannot be databased in a meaningful way to create systematized knowledge and a science.

A second drawback is the inability to detect interactions that are unsuspected or unknown. Most approaches using conjoint analysis test only a limited number of preselected combinations of elements (features of the products). Special combinations of elements must be added to detect interactions between pairs of elements, so that knowledge of such interactions must exist ahead of time and then such knowledge will dictate the precise combination to test. Overcoming these two limitations could improve the NPD process.

Experimental design is a key to absolute utilities on the one hand and discovering interactions between the elements (product features) on the other. The approach to solving the two problems is based on IPEDs, which includes all possible pairwise combinations of elements, multiple times across many respondents. The subsequent regression analysis, done either at the individual respondent level or at the group level, generates absolute values of the utilities. The regression analysis done at the group level partials out the linear effects and then allows for pairwise interactions to be detected and quantified. With enough permutations from a single design the approach can detect interactions among triples of elements and possibly higher interactions as well.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

DISCLOSURE

This chapter is partially based on the author's doctoral dissertation: Gofman, A. (2009). Experimentation-based product development in mature food categories: Advancing conjoint analysis approach. Tartu, Estonia: Tartu University Press.

REFERENCES

Aiken, L., & West, S. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage Publications.

Anderson, N. (1977). Functional measurement and psychological judgment. *Psychological Review, 77,* 153–170.

Atkinson, A., & Bailey, R. (2001). One hundred years of the design of experiments on and off the pages of biometrika. *Biometrika, 88*(1), 53–97.

Atkinson, A., & Donev, A. (1992). *Optimum experimental designs*. Oxford, UK: Clarendon Press.

Ayers, D., Dahlstrom, R., & Skinner, S. (1997). An exploratory investigation of organizational antecedents to new product success. *Journal of Marketing, 34*(1), 107–116.

Balasubramanian, S., Krishnan, V., & Sawhney, M. (2001). New offering realization in the networked digital environment. In J. Wind & V. Mahajan (Eds.), *Digital marketing* (pp. 310–338). New York, NY: John Wiley and Sons, Inc.

Beckley, J., Foley, M., Topp, E., Huang, J., & Prinyawiwatkul, W. (2007). *Accelerating new food product design and development (Institute of Food Technologists Series)*. Hoboken, NJ: Wiley-Blackwell.

Belliveau, P., & Griffin, A. (2002). *PDMA new product development tool book*. Hoboken, NJ: John Wiley and Sons, Inc.

Belson, K. (2003, February 2). Sony again turns to design to lift electronics. *New York Times*. Retrieved December 4, 2009, from http://www.nytimes.com/2003/02/02/business/yourmoney/02SONY.html?pagewanted=4#

Berghman, L., Matthyssens, P., & Vandenbempt, K. (2006). Building competences for new customer value creation: An exploratory study. *Industrial Marketing Management, 35,* 961–973.

Bretscher, O. (1995). *Linear algebra with applications* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Buisson, D. (1995). Developing new products for the consumer. In D. Marshall (Ed.), *Food choice and the consumer* (pp. 182–215). Cambridge, MA: Chapman and Hall.

Calantone, R., Cavuşgil, S. T., & Zhao, Y. (2002). Learning orientation, firm innovation capability and firm performance. *Industrial Marketing Management, 31*(6), 515.

Chrzan, K., & Orme, B. (2000). An overview and comparison of design strategies for choice-based conjoint analysis. In Sawtooth Software Conference (Ed.), *Proceedings of 2000 Sawtooth Software Conference* (pp. 161–177). Sequim, WA: Sawtooth Software.

Dahan, E., & Mendelson, H. (1998, October). *Optimal parallel and sequential prototyping in new product design*. Working paper. Stanford, CA: Stanford University Graduate School of Business, 1–53.

Dahan, E., & Mendelson, H. (2001). An extreme-value model of concept testing. *Management Science (Design and Development Issue), 47*(1), 102–116.

Davenport, T., & Klahr, P. (1998). Managing customer support knowledge. *California Management Review, 40*(3), 195–208.

De Luca, L., & Atuahene-Gima, K. (2007). Market knowledge dimensions and cross-functional collaboration: Examining the different routes to product innovation performance. *Journal of Marketing, 71,* 95–112.

Deshpandé, R., & Zaltman, G. (1982). Factors affecting the use of market research information: A path analysis. *Journal of Marketing Research, 19*(1), 14–31.

Drucker, P. F. (1995). *Managing in a time of great change*. New York, NY: Truman Talley Books/Dutton.

Evgenoiu, T., Boussios, C., & Zacharia, G. (2002). *Generalized robust conjoint estimation.* Working paper. Fontainebleau, France: INSEAD.

Flavin, R. (2008, May 4). *Top Ten New Product Launch Mistakes*. Retrieved July 27, 2009, from http://ezinearticles.com/?id=1152926.

Flores, L., Moskowitz, H. R., & Maier, A. (2003). From "weak signals" to successful product development: Using advanced research technology for consumer driven innovation. In ESOMAR (Ed.), *Proceedings of ESOMAR Conference Technovate* 2. Cannes: ESOMAR.

Fuller, G. (1994). *New food product development. From concept to marketplace*. Boca Raton, FL: CRC Press.

Gofman A. (2009). *Experimentation-based product development in mature food categories: Advancing conjoint analysis approach*. Tartu, Estonia: Tartu University Press.

Gold, A., Malhotra, A., & Segars, A. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185–214.

Goldman, E. (2005). Developing true product innovations: Coming up with innovative food products affects all aspects of a business and a new product initiative may need outside help to be successful. *Prepared Foods*. Retrieved October 5, 2008, from http://www.findarticles.com/p/articles/mi_m3289/is_10_174/ai_n15726041

Green, P. E., & Devita, M. (1975). An interaction model of consumer utility. *The Journal of Consumer Research, 2*(2), 146–153.

Green, P. E., & Krieger, A. M. (1991). Segmenting markets with conjoint analysis. *Journal of Marketing, 55,* 20–31.

Green, P. E., & Srinivasan, V. (1981). A general approach to product design optimization *via* conjoint analysis. *Journal of Marketing, 45,* 17–37.

Green, P. E., & Srinivasan, V. (1990). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5,* 103.

Green, P. E., & Wind, Y. (1973). *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, IL: Dryden Press.

Harrell, F. Jr. (2008). *Regression modeling strategies*. Nashville, TN: Vanderbilt University. Retrieved July 31, 2008, from http://biostat.mc.vanderbilt.edu/twiki/pub/Main/RmS/rms.pdf

Hauser, J., & Rao, V. (2003). Conjoint analysis, related modeling and applications. In P. Green & Y. Wind (Eds.), *Marketing research and modeling: Progress and prospects: A tribute to Paul Green* (pp. 141–168). New York, NY: Springer.

von Hippel, E. (1976). The dominant role of users in the scientific instrument innovation process. *Research Policy, 5,* 213–239.

von Hippel, E. (1978). Successful industrial products from customer ideas. *Journal of Marketing, 40*(1), 39–49.

von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science, 32*(7), 791–805.

von Hippel, E. (1988). *The sources of innovation*. New York, NY: Oxford University Press.

von Hippel, E. (2005). *Democratizing innovation*. Cambridge, MA: MIT Press.

von Hippel, E., & Katz, R. (2002). Shifting innovation to users *via* toolkits. *Management Science, 48*(7), 821–833.

Hair Jr, J. anderson, R., Tatham, R., & Black, W. (1995). *Multivariate data analysis: with readings*: Upper Saddle River, NJ, USA. Prentice-Hall, Inc.

Hunt, S. (2000). A general theory of competition: Resources, competences, productivity, economic growth: Sage Publications.

Hunt, S., & Morgan, R. (1995). The comparative advantage theory of competition. *Journal of Marketing, 59,* 1–15.

Hurley, R., & Hult, G. T. (1998). Innovation, market orientation and organizational learning: An integration and empirical examination. *Journal of Marketing, 62,* 42–54.

IBM. (2008). *Global CEO Study-IBM, The Enterprise of the Future*. Retrieved January 6, 2008, from http://www.935.ibm.com/services/us/gbs/bus/html/bcs_strategyplanning.html

Inkpen, A., & Dinur, A. (1998). Knowledge management processes and international ventures. *Organizational Science, 9*(4), 454–468.

Iyengar, V., Lee, J., & Campbell. M. (2001). Q-Eval: Evaluating multiple attribute items using queries. In *Proceedings of the Third ACM Conference on Electronic Commerce* (pp. 144–153). Tampa, FL: Association for Computing Machinery.

Jaeger, S., Rossiter, K., Wismer, W., & Harker, F. (2003). Consumer-driven product development in the kiwifruit industry. *Food Quality and Preference, 14,* 187–198.

Johnson, R. M. (1984). Trade-off analysis of consumer values. *Journal of Marketing Research, 11*(2), 121–127.

Kantowitz, B., Roediger, H., & Elmes, D. (2008). *Experimental psychology*. Florence, KY: Wadsworth Publishing.

Kiley, D. (2005, November 14). Shoot the focus group. *Business Week,* 120–121.

Klink, R., & Athaide, G. (2006). An illustration of potential sources of concept-test error. *Journal of Product Innovation Management, 23*(4), 359–370.

Krieger, A., Green, P.E, & Wind, Y. (2004). *Adventures in conjoint analysis: A practitioner's guide to trade-off modeling and applications*. Philadelphia, PA: Wharton School, University of Pennsylvania.

Kuhfeld, W., Tobias, R., & Garratt, M. (1994). Efficient experimental design with marketing research applications. *Journal of Marketing Research*, 31(4), 545-557.

Lees, G., & Wright, M. (2004). The effect of concept formulation on concept test scores. *Journal of Product Innovation Management, 21*(6), 389–400.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal, 13,* 111–125.

Li, T., & Calatone, R. (1998). The impact of market knowledge competence on new product advantage: Conceptualization and empirical examination. *Journal of Marketing, 62*(4), 13–29.

Lord, J. (1999). New product failure and success. In A. Brody & J. Lord (Eds.), *Developing new food products for a changing marketplace* (pp. 55–86). Lancaster, PA: Technomic Publishing Company.

Louviere, J. (1988). Conjoint analysis modelling of stated preferences: a review of theory, methods, recent developments and external validity. *Journal of Transport Economics and Policy*, 10: 93-119.

McDonough III, E., Kahn, K., & Barczak, G. (2001). An investigation of the use of global, virtual and colocated new product development teams. *The Journal of Product Innovation Management, 18,* 110–120.

Millett, S. (2006). Futuring and visioning: Complementary approaches to strategic decision making. *Strategy & Leadership, 34*(3), 43–50.

Moskowitz, H., & Gofman, A. (2007). *Selling blue elephants: How to make great products that people want BEFORE they even know they want them*. Upper Saddle River, NJ: Wharton School Publishing/Pearson Education.

Moskowitz, H. R., Porretta, S., & Silcher, M. (2005). *Concept research in food product design and development*. Ames, IA: Blackwell Professional Publishing.

Möller, K. (2006). Role of competences in creating customer value: A value-creation logic approach. *Industrial Marketing Management, 35,* 913–924.

Neal, W. D. (2003). Principles of market segmentation. *Marketing Power*. Retrieved May 17, 2008, from http://www.marketingpower.com/live/content.php?Item_ID=1006

Olson, E, Walker, O. Jr., & Ruekert, R. (1995). Organizing for effective new product development: The moderating role of product innovativeness. *Journal of Marketing, 59*(1), 48–62.

Orme, B. (2006). *Getting started with conjoint analysis: Strategies for product design and pricing research.* Madison, WI: Research Publishers LLC.

Pangborn, R. M. (1970). Individual variations in affective responses to taste stimuli. *Psychonomic Science, 21,* 125–128.

Rabino, S., Moskowitz, H., Katz, R., Maier, A., Paulus, K., Aarts, P., Beckley, J. and Ashman, H. 2007. Creating Databases from Cross-National Comparisons of Food Mind-Sets. *Journal of Sensory Studies*, 22(5): 550-586.

Rao, C., Toutenburg, H., Fieger, A., Heumann, C., Nittner, T., & Scheid, C. (1999). *Linear models: Least squares and alternatives, springer series in statistics*. New York, NY: Springer.

Riquelme, H. (2001). Do consumers know what they want? *Journal of Consumer Marketing, 18*(5), 437–448.

Rudolph, M. (1995). The food product development process. *British Food Journal, 97*(3). 3–37.

Ruggles, R. (1998). The state of the notion: Knowledge management in practice. *California Management Review, 40*(3), 80–89.

Sanchez, R., & Mahoney, J. (1996). Modularity, flexibility and knowledge management in product and organization design. *Strategic Management Journal, 17,* 63–76.

Sasena, M., Parkinson, M., Reed, M., Papalambros, P., & Goovaerts, P. (2005). Improving an ergonomics testing procedure *via* approximation-based adaptive experimental design. *Journal of Mechanical Design, 127*(5), 1006–1013.

Schlegelmich, B., & Penz, E. (2002). Knowledge management in marketing. *The Marketing Review, 3*(1), 5–19.

Sinkula, J., Baker, W., & Noordewier, T. (1997). A framework for market-based organizational learning: Linking values, knowledge and behavior. *Journal of the Academy of Marketing Science, 25,* 305–318.

Thomke, S. (1998). Managing experimentation in the design of new products. *Management Science, 44*(6), 743–762.

Thomke, S. (2001). Enlightened experimentation. The new imperative for innovation. *Harvard Business Review, 79*(2), 66–75.

Thomke, S. (2003). *Experimentation matters: Unlocking the potential of new technologies for innovation*. Boston, MA: Harvard Business School Press.

Toubia, O. (2004). *New approaches to idea generation and consumer input in the product development process*. Cambridge, MA: MIT Press.

Toubia, O., Simester, D., Hauser, J., & Dahan, E. (2003). Fast polyhedral adaptive conjoint estimation. *Marketing Science, 22*(3), 273–303.

Trott, P. (2008). *Innovation management and new product development.* Upper Saddle River, NJ: Prentice Hall.

Vriens, M. (1995). *Conjoint analysis in marketing. Developments in stimulus representation and segmentation methods* (Doctoral dissertation, Groningen University SOM). Capelle a/d Ijssel, Netherlands: Labyrint Publication.

Walter, J., Lechner, C., & Kellermanns, F. (2007). Knowledge transfer between and within alliance partners: Private *versus* collective benefits of social capital. *Journal of Business Research, 60*(7), 698–710.

CHAPTER 2

Isomorphic Permuted Experimental Designs in Conjoint Analysis

Alex Gofman and Howard R. Moskowitz*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The chapter deals with experimental designs used in conjoint analysis. The approach permutes the structure of the underlying fractional experimental design in order to create different sets of combinations. The resulting experimental designs, called isomorphic permuted experimental designs (IPEDs), create diverse sets of the variables and levels, producing an array of different designs that are statistically equivalent to each other. By creating an array of distinctive different individual designs (one design for each respondent), IPEDs reduce the bias caused by some possibly unusually strong performing combinations. IPEDs create the conditions for statistical analyses to detect and estimate interactions among variables. IPEDs also allow cluster analysis to identify pattern-based segments emerging from individual models of utilities. The chapter presents the theoretical foundation of the approach, formalizes the algorithmic implementation and shows a practical example its use.

Keywords: Conjoint analysis, experimental design, regression analysis, fractional experimental designs, individual designs, dummy variable regression, incomplete concepts, interactions, pattern-based segmentation.

INTRODUCTION

Conjoint analysis has become increasingly prevalent as a major approach to studying consumer preferences. Green and Srinivasan (1990a, 1990b), Green and Krieger (1991) and Krieger, Green and Wind (2004) demonstrate that conjoint analysis enjoys a number of practical advantages when used to quantify consumer preferences. Conjoint analysis begins by assuming that a product or service can be deconstructed into its component *variables* (also called *attributes, silos*, or *categories*) and *levels* (*elements*). By presenting to respondents a series of *profiles (concepts)*, *i.e.* combinations of levels from different variables and getting their subjective judgments, conjoint analysis ends up uncovering the utility value for each level of each stimulus variable. These are called the individual utilities, or equivalent, the part-worth contributions, or impact scores, respectively.

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Experimental design, a statistical plan that lays out the combinations of the profile elements*,* lies at the heart of conjoint analysis, as it does for numerous other experiment-based approaches to building knowledge. Conjoint analysis has enjoyed the attention of many researchers. Despite the substantial literature of research already published and available in many works (*e.g.,* Cattin and Wittink, 1982; Carroll and Green, 1995; Atkinson and Haines, 1996; and Atkinson and Bailey, 2001, to name a few), the field of conjoint analysis is still active and presents an opportunity for research that will at once be important in theory and in practice.

The terminology of conjoint analysis, in particular and of experimental design in general, can be confusing. We will adopt the nomenclature offered by one of the leaders in the field of experimental design, George Box. Box, Hunter and Hunter (2005) summarize and contrast a variety of experimental designs differing in the number of *variables* (*factors*; applied to attributes in conjoint analysis) and the number of *levels* (matched to the elements in conjoint analysis). These variables and levels are then combined into *runs* (*experimental units* or rows of design), which are the same as profiles in the language of conjoint analysis. To keep matters consistent and to make the terminology a friend rather than a complicating foe, we will use terms "variables", "levels" and "profiles", applying those terms to experimental design in general and to conjoint analysis in particular.

As in any practical method for acquiring and analyzing data, conjoint analysis has strengths but it also has some rather severe limitations. Moskowitz, Porretta and Silcher (2005) point to the limitations of some traditional experimental designs. The first limitation is *scope*. In traditional approaches, the experimental design creates only one set of test stimuli. The design approach is applied to a set of variables only once creating a single set of combinations for all respondents. Of course, the specifics combinations in the design can be tested in a random order, to reduce some of the order bias. The second limitation is more insidious, its *inescapable collinearity*. The traditional approaches often create the test stimuli by creating *complete concepts*. This is the case for so-called full-profile conjoint analysis, in which every test stimulus, *i.e.* every combination of the levels, must have one and only one level of each of the variables present in the stimulus. In such cases, one cannot estimate the absolute utility value of a level. Rather, the utility values are estimated relative to a *reference level*—one of the subjectively selected levels. With the complete concepts approach, one cannot compare the utilities of levels across different variables. Rather, one can only compare the utilities of levels within the same variable.

Biases and limitations plague the traditional methods of constructing and testing the stimuli in conjoint analysis. Moskowitz *et al.* (2005) and Gofman (2006) summarize the limitations of the existing approaches and argue that they create severe interlinked statistical problems:

- *Biased results.* A limited number of distinct prototypes lead to a bias in outcome. A few potentially very strong levels might skew the results.

- *Insufficient variation of combinations* also prevents detection and estimation of interactions (pair wise and higher order).

- *Collinearity*. Complete concepts create multicollinearity, which prevents the analysis from estimating the true utility value of profile levels.

- *No true estimate of the basic level of interest*. The statistical analyses of such complete concepts require effects model regressions, in which there is no estimate of the additive constant (the basic level of consumer interest) and the requirement that the utilities of the levels in each variable add up to zero.

- *No true estimate of the utility value of the individual levels*. The requirement of the constant sum equal to zero means that when a new level is brought into the study, the utilities of the other levels must be readjusted because they have relative value. This readjustment means that one cannot use the results for databasing. The absolute values of the utilities, having no real meaning and being susceptible to the effects of other variables in the study, cannot be compared across variables or projects unless precisely the same variables are used from study to study.

Moskowitz (1994) suggested a practical approach that permutes the structure of the underlying fractional design to make multiple different sets of combinations.

However, in that early publication there was no generalized model/description of the approach. Initial steps toward formalization of the process were described in Moskowitz and Gofman (2005) and Gofman (2006). This chapter further develops the permutation strategy applied to experimental designs.

PERMUTING EXPERIMENTAL DESIGNS

An alternative to the complete concepts approach was presented by Moskowitz *et al.* (2005). The approach was entitled *incomplete concepts* (with profiles having *zero conditions*). Arraying the combinations of levels in a specified experimental design with true zeros, *i.e.* with some combinations entirely missing a variable, enables the research to estimate the absolute values of the utilities. These designs require more profiles, however, than other designs without such zero conditions, for the same degree of balancing, where each of the levels appears equally often. The slight inconvenience of having more profiles is, in most cases, compensated by the ability of these designs to generate absolute values. Those absolute values enable comparisons between variables in the same study and databasing of the variables across time and studies with different arrays of elements. As such, the strategy constitutes the first stage in building scientific knowledge based upon a large collection of variables and their associated utilities.

Let us analyze the approach based on a particular fractional experimental design, the Plackett Burman 5-Level screening design (Encyclopedia of Statistical Sciences, 1985) shown in Table **1**. This particular design enables the researcher to investigate up to five variables in a profile in conjoint analysis and up to four levels per variable. Note that whereas the experimental design allows for five levels per variable, the fifth level is reserved for “null” or “no level present” (signified by “0” in Table **1**). By allowing for a true “zero condition”, the researcher can use the regression analysis to better estimate the contribution of every level to respondent reactions Moskowitz *et al.* (2005).

Let us now see how this new approach fits into the world of experimental design. A conceptual model of the traditional approach to experiential design appears in the left part of Fig. **1**. The experimental design is applied to a set of variables only one time. That single application creates one design to be tested by all of the

Table 1: An Example of an Experimental Design for 25 Profiles, Based on the Plackett Burman Screening Design.

N	Design				
	Var[1]	**Var**[2]	**Var**[3]	**Var**[4]	**Var**[5]
1	x_4^1	x_1^2	x_3^3	x_1^4	x_1^5
2	x_0^1	x_4^2	x_1^3	x_3^4	x_1^5
3	x_3^1	x_0^2	x_4^3	x_1^4	x_3^5
4	x_3^1	x_3^2	x_0^3	x_4^4	x_1^5
5	x_2^1	x_3^2	x_3^3	x_0^4	x_4^5
6	x_3^1	x_2^2	x_3^3	x_3^4	x_0^5
7	x_4^1	x_3^2	x_2^3	x_3^4	x_3^5
8	x_1^1	x_4^2	x_3^3	x_2^4	x_3^5
9	x_2^1	x_1^2	x_4^3	x_3^4	x_2^5
10	x_2^1	x_2^2	x_1^3	x_4^4	x_3^5
11	x_0^1	x_2^2	x_2^3	x_1^4	x_4^5
12	x_2^1	x_0^2	x_2^3	x_2^4	x_1^5
13	x_4^1	x_2^2	x_0^3	x_2^4	x_2^5

N	Design (cont.)				
	Var[1]	**Var**[2]	**Var**[3]	**Var**[4]	**Var**[5]
14	x_3^1	x_4^2	x_2^3	x_0^4	x_2^5
15	x_0^1	x_3^2	x_4^3	x_2^4	x_0^5
16	x_0^1	x_0^2	x_3^3	x_4^4	x_2^5
17	x_1^1	x_0^2	x_0^3	x_3^4	x_4^5
18	x_0^1	x_1^2	x_0^3	x_0^4	x_3^5
19	x_4^1	x_0^2	x_1^3	x_0^4	x_0^5
20	x_2^1	x_4^2	x_0^3	x_1^4	x_0^5
21	x_1^1	x_2^2	x_4^3	x_0^4	x_1^5
22	x_1^1	x_1^2	x_2^3	x_4^4	x_0^5
23	x_3^1	x_1^2	x_1^3	x_2^4	x_4^5
24	x_1^1	x_3^2	x_1^3	x_1^4	x_2^5
25	x_4^1	x_4^2	x_4^3	x_4^4	x_4^5

Note. $\boldsymbol{x_j^i}$ is a permuted design experimental unit for variable *i* and level *j* (0 denotes a missing level).

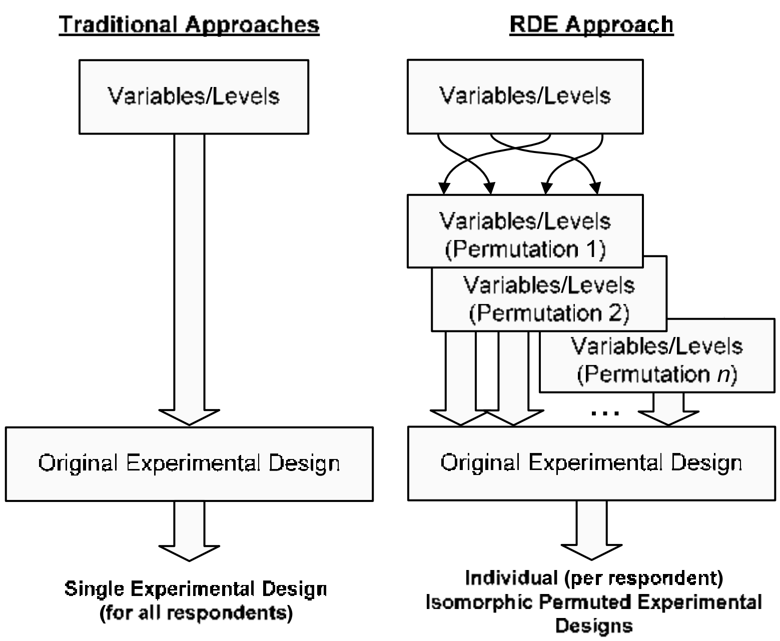


Figure 1: Conceptual model of traditional experimental designs approaches (on the left) and proposed individual isomorphic permuted experimental designs (IPEDs on the right).

respondents. In practice, the different profiles that are created by this one application of the design will be randomized to reduce order bias, but the randomization will always be made using the same set of combinations. A profound drawback of this single application is that the experimental design creates only a limited number of combinations. One cannot detect all interactions, nor indeed even a significant proportion of the combination. There is also a possibility of a bias introduced due to the limited and fixed number of concepts tested by all the respondents Moskowitz *et al.* (2005).

When we randomize the sequence of the variables and levels inside the variables *before* applying the experimental design, we create a large number of individual designs comprising unique concepts up to statistical limits imposed by the specific design. The randomization, *i.e.* permutation, approach creates *isomorphic designs*, which are statistically equivalent to each other. The strategy is called isomorphic permuted experimental designs (IPEDs).

The conceptual model of IPED (Fig. **1**, right) shows multiple permutations of the variables/levels. The permutations occur *before* the experimental design is used to create the combinations. The permutation strategy creates a distinct design for each respondent. At the group level, with all of the combinations considered as one group, there are thousands of unique combinations. With sufficient number of individually permuted designs all of the pairwise combinations of elements from different variables occur in the set. This permutation strategy creates a more bias-free environment to test the variables and levels because unexpected interactions among pairs of elements are averaged out (Moskowitz *et al.* 2005). In addition, the permutation strategy creates a database of information that can explicitly identify the utility value of all pairwise interactions of elements coming from different variables (Gofman, 2006).

A DEMONSTRATION OF THE APPROACH

According to the steps described in Gofman and Moskowitz (2009), the following demonstrates the approach using the example of concepts for donuts, a mature food category used for snacks. For the purposes of our demonstration, the sensory, image, usage and other descriptions of donuts are structured as a set of four variables with three levels in each (Table **2**). This structuring is done according to the predilections of the researchers.

Table 2: Variables and Levels of the Demonstration Project on Donuts.

Code	Elements
Variable A: Benefit	
A1	Simply the best cinnamon rolls in the whole wide world
A2	Made fresh … especially for you … by you
A3	From your favorite local bakery or pastry shop
Variable B: Emotional	
B1	A joy for your senses … seeing, smelling, tasting
B2	It feeds THE HUNGER
B3	When you think about it, you have to have it…and after you have it, you can't stop eating it
Variable C: Primary attribute	
C1	Big, three-inch spiraled rounds of dense chewy pastry like a donut with sweet cinnamon inside, covered with sweet icing
C2	Huge, thick, four-inch spiraled rounds of light flaky pastry with sweet cinnamon inside, covered in a cream cheese frosting
C3	The ultimate chocolate indulgence with rich chocolate inside a huge, thick and gooey spiraled cinnamon bun with sweet icing and a gooey chocolate dripping over the top
Variable D: Mood	
D1	Premium quality…that great classic taste, like it used to be
D2	With extra chocolate, cream cheese, or sugary icing on the side just waiting for dipping
D3	100% natural…and new choices every month to keep you tantalized

The study utilizes the Plackett Burman 4-Factor 4-Level screening design (Encyclopedia of Statistical Sciences, 1985) with one level in each factor reserved for "zero condition". This fractional design requires 20 profiles for each respondent (Table **3**). Here, x_0^i represents "zero condition" for category *i* when the category is absent from the test profile. Each level x_j^i (experimental unit for factor *i*, variable *j*) is applied to a set of variables and levels from a specific project. In our case, it would be an individually permuted selection of variables and levels.

The process results in an individual experimental design for each respondent that is unique yet isomorphic. Table **4** shows examples of two permuted individual experimental designs, one design for each of two respondents. This table also shows the design in a form prepared for dummy variable regression, following the conventions of the typical off-the-shelf software system for statistical analysis. In

Table **4**, levels of the variable e^i_j (level *j* of variable *i*) take the value "1" when they are present in a concept (arrayed as a row) and take on the value "0" when they are absent. Together, the levels comprise the independent variables in regression and the rating comprises the dependent variable (see the last column in Table **4**).

Table 3: Example of Original (Source) Experimental Design Used as the Basis for the Subsequent Individual Permutations.

Design				
Unit	**Var¹**	**Var²**	**Var³**	**Var⁴**
1	x^1_3	x^2_2	x^3_0	x^4_3
2	x^1_3	x^2_0	x^3_0	x^4_2
3	x^1_2	x^2_0	x^3_3	x^4_3
4	x^1_2	x^2_0	x^3_2	x^4_2
5	x^1_0	x^2_3	x^3_3	x^4_3
6	x^1_0	x^2_2	x^3_2	x^4_1
7	x^1_3	x^2_3	x^3_3	x^4_1
8	x^1_0	x^2_2	x^3_1	x^4_1
9	x^1_1	x^2_3	x^3_1	x^4_2
10	x^1_0	x^2_1	x^3_1	x^4_0

Design (cont.)				
Unit	**Var¹**	**Var²**	**Var³**	**Var⁴**
11	x^1_1	x^2_1	x^3_2	x^4_3
12	x^1_3	x^2_1	x^3_0	x^4_2
13	x^1_1	x^2_2	x^3_3	x^4_0
14	x^1_1	x^2_0	x^3_2	x^4_3
15	x^1_2	x^2_3	x^3_0	x^4_1
16	x^1_2	x^2_2	x^3_3	x^4_2
17	x^1_3	x^2_0	x^3_1	x^4_0
18	x^1_0	x^2_3	x^3_2	x^4_0
19	x^1_2	x^2_1	x^3_0	x^4_0
20	x^1_1	x^2_1	x^3_1	x^4_1

Table 4: Examples of Two Permuted Individual Experimental Designs (for Two Respondents) with Ratings Assigned by Respondents to Each Individual Profile.

Unit	**Levels**												**Rating**
	e^1_1	e^1_2	e^1_3	e^2_1	e^2_2	e^2_3	e^3_1	e^3_2	e^3_3	e^4_1	e^4_2	e^4_3	
Respondent 1													
1	0	0	0	0	1	0	0	0	1	1	0	0	9
2	0	0	1	0	0	1	0	0	0	1	0	0	5
3	0	0	1	1	0	0	0	0	1	0	0	0	5
4	1	0	0	1	0	0	1	0	0	0	0	0	5
5	0	1	0	0	1	0	0	1	0	0	1	0	5
6	0	1	0	0	0	0	1	0	0	1	0	0	5
7	0	0	1	0	0	0	0	0	1	0	0	1	5
8	0	0	0	0	0	0	0	1	0	0	1	0	9
9	0	1	0	0	0	0	0	1	0	1	0	0	9
10	0	1	0	0	0	1	0	0	1	0	0	1	7
11	0	0	1	0	1	0	1	0	0	0	0	0	5

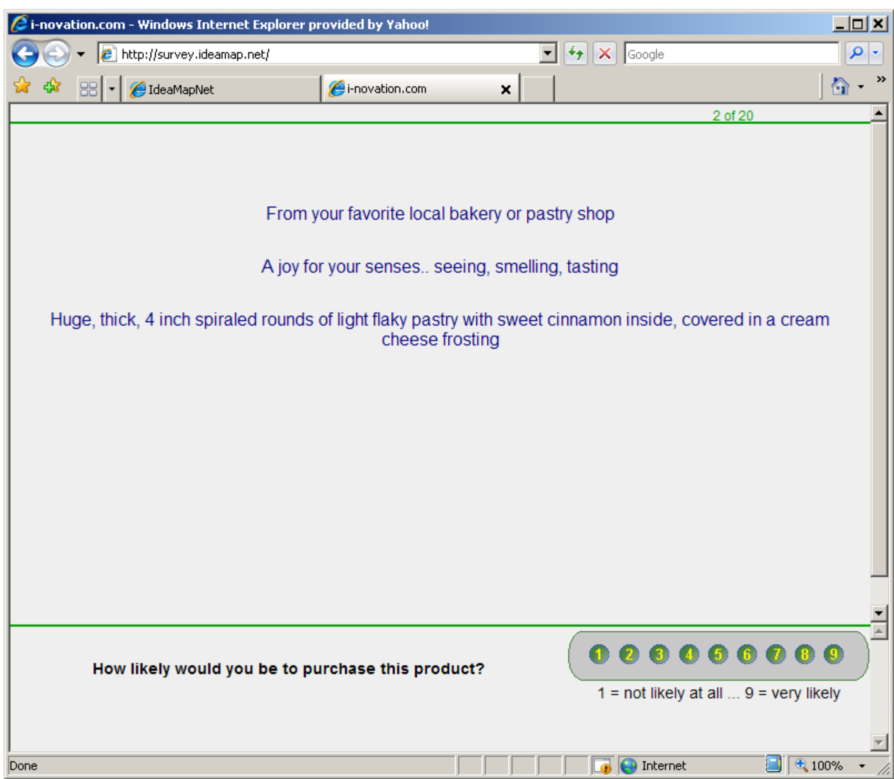
Unit	**Levels**												**Rating**
	e^1_1	e^1_2	e^1_3	e^2_1	e^2_2	e^2_3	e^3_1	e^3_2	e^3_3	e^4_1	e^4_2	e^4_3	
Respondent 2													
1	1	0	0	0	0	1	0	0	0	0	1	0	1
2	0	0	1	1	0	0	0	0	0	0	0	0	4
3	0	0	0	1	0	0	0	0	0	1	0	0	1
4	0	0	0	0	0	1	1	0	0	0	0	1	1
5	0	0	1	0	0	0	0	1	0	0	1	0	3
6	0	1	0	0	0	0	0	0	1	0	0	0	7
7	0	1	0	0	1	0	1	0	0	0	1	0	2
8	0	0	0	0	0	1	1	0	0	1	0	0	1
9	0	1	0	1	0	0	0	0	1	0	0	0	1
10	0	0	1	0	0	1	0	0	1	0	1	0	1
11	1	0	0	1	0	0	0	1	0	0	0	1	1

Table 4: cont….

15	0	0	0	1	0	0	0	0	0	0	1	0	7	15	0	0	1	0	1	0	1	0	0	0	0	0	1
16	1	0	0	0	1	0	0	1	0	0	0	1	2	16	1	0	0	1	0	0	1	0	0	1	0	0	1
17	1	0	0	1	0	0	0	0	1	1	0	0	2	17	0	1	0	0	0	0	0	0	0	1	0	0	7
18	0	1	0	1	0	0	0	0	0	0	0	1	6	18	0	1	0	0	0	1	0	1	0	0	0	0	1
19	1	0	0	0	0	1	0	0	0	0	0	0	3	19	0	0	0	0	1	0	0	1	0	0	1	0	1
20	0	0	0	0	0	0	1	0	0	0	0	1	9	20	1	0	0	0	1	0	0	0	1	1	0	0	1

Note. The permuted designs are based on the source design (Table **3**).

Fig. **2** shows sample screen captures of the interview in which a respondent evaluates experimentally designed concepts on a 1 to 9 rating scale. Every respondent sees a set of unique (up to the statistical limits of the specific design) concepts. In the example above, the respondents evaluate 20 screens each. As can be seen, some concepts have four levels present while others have "zero conditions"—intentionally missed levels.



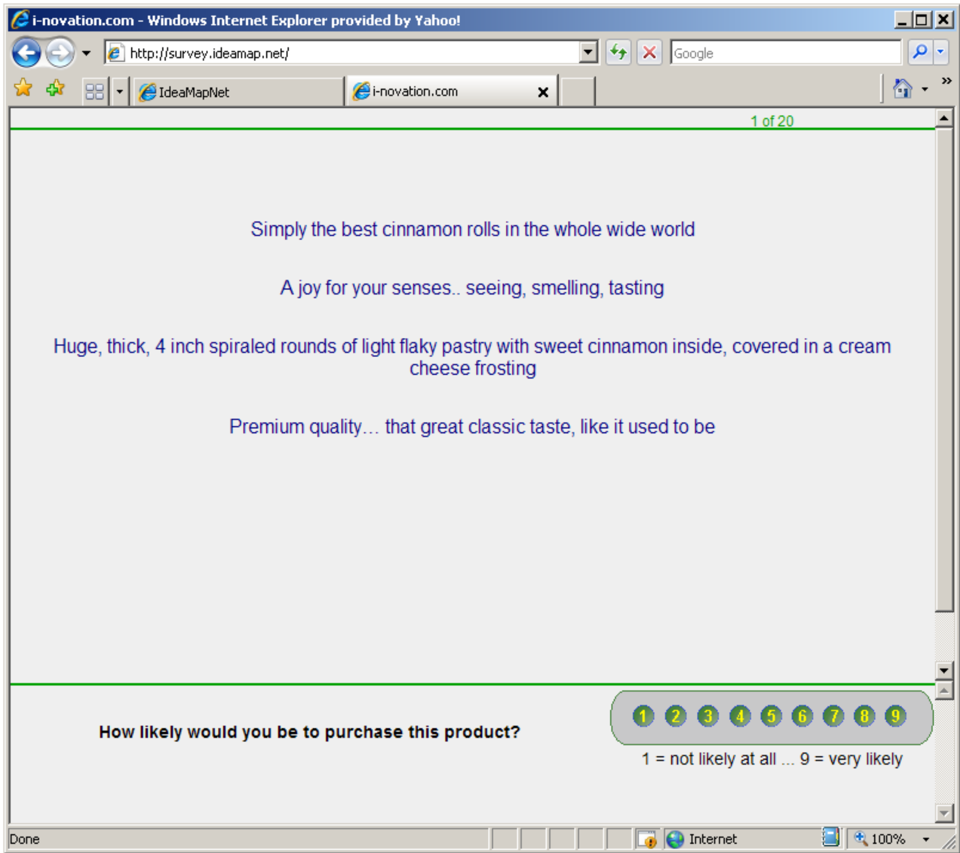


Figure 2: Sample screen captures of a respondent interview (utilizing the IdeaMap®.NET online tool). The first profile has one variable missing.

The detailed results interpretation, pattern-based segmentation of consumers and interaction analyses can be found in Gofman and Moskowitz (2009) and Gofman (2006).

CONCLUSIONS

The permuted designs establish a new capability to detect all pairwise interactions between levels of different attributes. With sufficient number of combinations tested, *i.e.* many respondents and many permuted designs, even higher-order interactions may be scrutinized and assessed, although their number is so great that the demands will be for an order of magnitude as there are more permuted designs and combinations to test. Nonetheless, the permutation strategy puts the evaluation of interactions and the discovery of synergisms and suppressions within the realm of possibility (Moskowitz and Gofman, 2005; Gofman, 2006).

For practical application, using IPEDs to create individual-level designs provides a set of benefits for conjoint analysis that bring the analysis forward. The benefits to conjoint analysis come from one of the two strategies. One strategy creates the set of designs and then analyzes each at the individual level (each person generates his own equation using regression). The second strategy uses the IPED as the source of different combinations, but groups the individual designs into one mass to be analyzed or at the group level (all individuals are combined into a dataset). The strategies considered together provide these four benefits:

1. Eliminate selection bias by creating many combinations, not just one limited set.

2. Test each pairwise combination among a reasonably large number of respondents without having to worry about setting up these combinations ahead of time.

3. Detect the existence of and estimate the magnitude of, the pairwise and possibly higher-order interactions between levels.

4. Cluster respondents based on the patterns of their utilities.

At the same time, we should be cognizant of some of the limitations. These are not necessarily drawbacks as much as they are practical considerations.

1. An average respondent has an attention span of about 15 min (Moskowitz *et al.* 2005) and thus could evaluate up to 60–75 concepts (approximately up to 40 levels).

2. For this size of conjoint analysis project, a sample of 200 or more respondents would produce a dataset for statistically significant analyses of pairwise interactions between the levels.

3. To detect higher-order interactions requires an order of magnitude more respondents.

4. In the majority of practical applications, the projects are executed without restrictions to avoid the complexity of setting and satisfying the constraints. As a consequence, most studies do not look for pairwise interactions, simply because there are not sufficient numbers of respondents. When interactions are important, one needs more generated, more permuted designs, which in turn invite more respondents to participate.

To conclude this chapter, let us look at the future, both in terms of the big picture of what the science can evolve to and what the discipline of consumer research can become. Any new advances in experimental design will soon find themselves embedded in different applications that use systematic experimentations. IPED-based work is no different. There are many new applications of the approach to the emerging areas such as Web page design and package optimization (Gofman, 2007; Gofman, Moskowitz and Mets, 2009a, 2010).

The advantages of the approach, such as the ability to database and compare results, opens the possibility of a new science, called Mind Genomics®, with the goal of identifying the pattern of segment membership in a particular individual for an array of behaviorally and economically meaningful products (Moskowitz, Gofman, Beckley and Ashman, 2006; Moskowitz, Gofman and Beckley, 2006; Gofman, Moskowitz and Mets, 2009b). Mind Genomics® spans the range between consumer research, psychology, genetics and economics.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This chapter is based on Gofman, A. and Moskowitz, H. (2010). Application of isomorphic permuted experimental designs in conjoint analysis. *Journal of Sensory Studies*. doi:10.1111/j.1745-459X.2009.00258.x

DISCLOSURE

Part of the information included in this chapter was previously published in *Journal of Sensory Studies* 25, Issue 1 (2010), 127–145. February 2010 (http://onlinelibrary.Wiley and Sons.com/doi/10.1111/j.1745-459X.2009.00258.x/abstract;jsessionid=2E2FD18FD06AB4D362D2E9DEF9AE91DE.d01t02)".

REFERENCES

Atkinson, A., & Bailey, R. (2001). One hundred years of the design of experiments on and off the pages of biometrika. *Biometrika, 88*(1), 53–97.

Atkinson, A., & Haines, L. (1996). Designs for nonlinear and generalized liner models. In R. Ghosh, & C. Rao (Eds.), *Handbook of statistics: Design and analysis of experiments* (pp. 437–475). Amsterdam: Elsevier.

Box, G., Hunter, W., & Hunter, J. (2005). *Statistics for experimenters: Design, innovation and discovery* (2nd ed.). New York, NY: Wiley and Sons, Inc.

Carroll, D., & Green, P. (1995). Guest editorial: Psychometric methods in marketing research: Part I, conjoint analysis. *Journal of Marketing Research, 32*(4), 385–391.

Cattin, P., & Wittink, D. R. (1982). Commercial use of conjoint analysis: A survey. *Journal of Marketing, 46,* 44–53.

Encyclopedia of Statistical Sciences. (1985). Multivariate analysis—Plackett and Burman designs. In S. Kotz, N. Johnson, & C. Read (Eds.), *Encyclopedia of statistical sciences*, Vol. 6. New York, NY: Wiley and Sons, Inc.

Gofman, A. (2006). Emergent scenarios, synergies and suppressions uncovered within conjoint analysis. *Journal of Sensory Studies, 21*(4), 373–414.

Gofman, A. (2007). Consumer driven multivariate landing page optimization: Overview, issues and outlook. *The IPSI BgD Transactions on Internet Research, 3*(2), 7–9.

Gofman, A., & Moskowitz, H. (2009). Steps towards a consumer-driven innovation machine for 'ordinary' product categories in their later lifecycle stages. *International Journal of Technology Management, 45*(3/4), 349–363.

Gofman, A., Moskowitz, H., & Mets, T. (2009a). Integrating science into web design: Consumer driven website optimization. *The Journal of Consumer Marketing, 26*(4), 286–298.

Gofman, A., Moskowitz, H., & Mets, T. (2009b). Developing new corporate understanding of an existing product. *The Journal of Product & Brand Management, 18*(2), 84–94.

Gofman, A., Moskowitz, H., & Mets, T. (2010). Accelerating structured consumer-driven package design. *The Journal of Consumer Marketing*, *27(*2), 157-168.

Green, P. E., & Srinivasan, V. (1990a). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5,* 103–123.

Green, P. E., & Srinivasan, S. (1990b). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing, 54*(4)*,* 3-19. Retrieved September 4, 2008, from http://knowledge.wharton.upenn.edu/chapters/775.pdf

Green, P. E., & Krieger, A. M. (1991). Segmenting markets with conjoint analysis. *Journal of Marketing, 55,* 20–31.

Krieger, A., Green, P. E, & Wind, Y. (2004). *Adventures in conjoint analysis: A practitioner's guide to trade-off modeling and applications*. Philadelphia, PA: Wharton School, University of Pennsylvania.

Moskowitz, H. R. (1994). A new paradigm for development: Rapid concept and product development in an age of resource scarcity. *Product Development & Management Association,* 66–85.

Moskowitz, H. R., & Gofman, A. (2005). U.S. Patent Application Serial No. 11/032,834 System and method for performing conjoint analysis. U.S. Patent Application Publication No. 2005/0177388.

Moskowitz, H., Gofman, A., & Beckley, J. (2006). Using high-level consumer-research methods to create a tool-driven guidebook and database for product development and marketing. *Journal of Sensory Studies, 21*(1), 54–100.

Moskowitz, H., Gofman, A., Beckley, J., & Ashman, H. (2006). Founding a new science: Mind Genomics®. *Journal of Sensory Studies, 21*(3), 266–307.

Moskowitz, H. R., Porretta, S., & Silcher, M. (2005). *Concept research in food product design and development*. Ames, IA: Blackwell Professional Publishing.

CHAPTER 3

Detecting Explicit and Implicit Interactions within Rule Developing Experimentation

Alex Gofman and Howard R. Moskowitz*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The chapter introduces two approaches that identify the nature and magnitude of interaction between concept elements in a conjoint analysis task. Both approaches use main effects experimental designs, permuted to create hundreds of new designs that are isomorphic to the original design structure. In the first approach, the scenario analysis creates a distinct, mutually exclusive, exhaustive set of subgroups from concepts based upon the commonality of a specific element, runs a dummy variable regression within each subgroup and identifies the effect of the different elements on the dependent variable. When compared across the different subgroups in the regression analysis, the outcome shows the effect of one element on the impact values of the other elements. In the second approach, also using regression analysis, this time to understand the pairwise interactions, the analysis forces in all of the linear terms (single elements) and then allows significant pairwise combinations to enter if they contribute significant additional predictability to the model. The two approaches identify the existence of and then measure the impact of, one element on the performance of others (scenario) and the unexpected effect of mixing two concept elements (interaction analysis). We illustrate the approaches with a case history dealing with communicating the sensory and refreshment benefits of an orange beverage.

Keywords: Conjoint analysis, interactions, synergism, suppression, experimental design, scenario analysis, regression analysis.

INTRODUCTION

The ability to discover just how pairs of communication elements interact with each other to drive consumer interest constitutes a major contribution to consumer research (Carmone, Green and Jain, 1978; Green and Srinivasan, 1990; Wittink and Cattin, 1989). When one knows that two elements synergize with each other so that their net effect is much larger than might be expected from measuring each alone, there is a clear opportunity to increase acceptance by putting these

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

synergistic elements together. Conversely, when two elements in combination work exceptionally poorly, it is good to know this unhappy fact as early as possible in the development process in order to ensure that these suppressive combinations simply are not allowed to appear. It makes most sense to benefit from synergies and avoid suppressions (Wittink, Vriens and Burhenne, 1994).

The question about interactions is simply "how" to make these discoveries in a statistically defensible way, in an efficient and meaningful way. From the consumer researcher's viewpoint, the existence and nature of these synergistic and suppressive mixtures often become points of mere conjecture. It is almost impossible to estimate how a combination of elements will perform when presented together in a single concept. It is often said by people in so-called creative roles in corporations and agencies that they are aware of these synergisms, that they just happen to know them. However, no clear evidence about this insight into synergism appears to have been published.

The term *interaction* may be defined in one of two different ways:

Implicit interactions (scenarios): the nature of responses when these responses are made within a certain set of conditions. In this first definition, we use the word "interaction" in a non-statistical way. Interaction is simply the effect of the conditions on some dependent variable. Thus, we might look for the interaction of a set of product features with breakfast by saying that knowing the consumer's mind-set of "breakfast", we look at how this mind-set affects reactions to test stimuli. Thus, interactions here really pertain to the effect of one variable on another. The magnitude of the interaction is the difference in responses to the same element when there are no conditions (*e.g.,* no mention of meal occasion) *versus* when there are conditions (*e.g.,* mention of a specific meal occasion). A good word for this type of interaction is "scenario". The conditions set up the scenario against which the response is provided.

Fig. **1** shows a conceptual model of implicit interactions. An attribute A_p selected as a pivotal one (*e.g.,* brands, key features) drives the utilities of the rest of the attributes. When the pivotal attribute A has m elements, there are m possible scenarios. In a conjoint analysis with total of N elements, for an element A_i, the scenario will analyze the utilities of the elements x_j^{pi} (where $j = 1\ldots n$; $n = N - m$) creating an n-component utilities array X^{pi}:

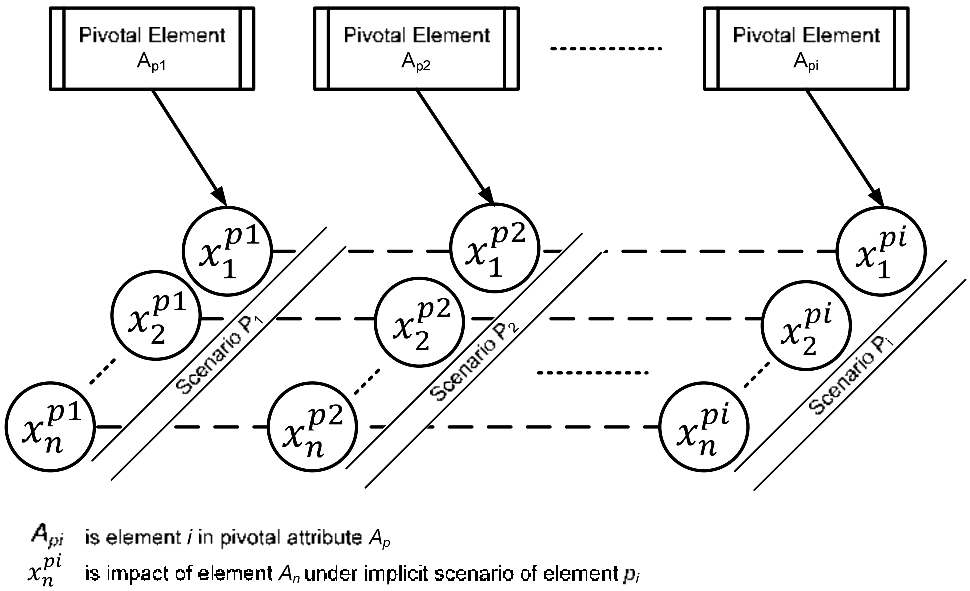


Figure 1: Conceptual model of implicit interactions (source: Gofman, 2009).

$$\boldsymbol{X^{pi}} = (\boldsymbol{x_1^{pi}}, \boldsymbol{x_2^{pi}}, \ldots, \boldsymbol{x_n^{pi}}).$$

Selection of another pivotal element will lead to a new, distinct scenario with a unique array of utilities for this implicit interaction. By comparing the performance of utilities under different pivoting elements, a researcher ends up with valuable input in the process of selecting features for NPD. The analysis shows which elements "work" with other elements.

The process of estimating the implicit interactions differs from the one process below for the explicit interactions of specific pairs of elements. If the attribute A is selected as a pivotal one, we can choose any of the elements in that attribute and select only those specific cases where this element was present. We ignore the other cases when this element was absent.

By repeating the selection process for every element in the chosen attribute, the researcher splits the whole data set into several different layers, *i.e.* strata. For E

elements in an attribute (*e.g.,* attribute A), there will be (E + 1) layers—one per each element in the attribute and one for cases when the attribute is absent because the experimental design is set up so that some of the concepts are missing elements from the particular attribute.

Explicit interactions (synergisms, suppressions): pairs of elements whose combination generates a "statistically" significant term in a regression equation or in analysis of variance. When responses to both individual elements and their combinations can be measured, it becomes possible to assess whether or not the combination performs far better or far worse than one might expect on the basis of the performance of the single element alone. This is the way that the statistician ordinarily thinks about and describes "interactions". Generally, interactions manifest themselves quantitatively in analysis of variance or regression once the proper mathematical expression has been made. For example, in regression analysis (used here), we can create an equation, which relates the presence/absence of two variables (A, B) to a rating:

$$Rating = k_0 + k_1 A + k_2 B + k_3 (A \times B) \ .$$

The equation shows how to estimate the magnitude of the rating from knowing the values of two variables (A, B). We can multiply the values of A and B in order to create an interaction term $A \times B$, which we then also insert into the equation. Through regression modeling we determine whether or not the interaction term, $A \times B$, is statistically significant and at what level of significance. When the term $A \times B$ is significant, we say that variables A and B interact with each other. When the interaction is significant and the coefficient of that interaction k_3 is positive, we conclude that A and B act synergistically. That is, when A and B combine with each other we can add an additional value k_3 to their sum of k_0, k_1 and k_2. When the coefficient of the interaction term, k_3, is negative, we subtract the value k_3 from the sum we expect when A and B appear together. In that case, we find a suppression between A and B. The values k_0, k_1, k_2 and k_3 must all be determined by regression analysis with the proper data set.

The previous chapter on isomorphic permuted experimental designs (IPEDs) describes the experimental design that helps to detect interactions in rule developing experimentation (RDE). We now describe the “work-products” from those conjoint analyses studies that use IPEDs.

The *Persuasion Model* show the intensity or magnitude of acceptance (Moskowitz, Gofman and Beckley, 2006; Moskowitz, Porretta and Silcher, 2005). The persuasion model comes directly from the regression analysis. The independent variables are the elements (which take on the value 1 when present, 0 when absent). The dependent variable is the actual rating on the scale. The regression first returns with the additive constant of the persuasion scale, k_0. This parameter is purely estimated value, showing the number of rating scale points that might be expected to be assigned to the concept if no elements were present. The additive constant is clearly an estimated value as all test concepts comprise two to four elements. The additive constant is a good “baseline” of interest deduced from the average ratings of all of the concepts. In turn, the coefficients $k_1...k_{36}$ for the persuasion model show the expected number of rating points to be added or subtracted from the additive constant if the element were to be inserted into the concept. The total utility of the concept is defined as the sum of the additive constant and the component utility values of the elements chosen to be part of the concept.

The *Interest Model* deals with membership in the acceptance class (Moskowitz *et al.* 2006, 2005). The initial steps to create the interest model at the individual respondent level transform the ratings to a binary value, either 0 or 100 using a specific criterion. (One criterion might be that ratings of 1–6 are transformed to 0; ratings 7–9 are transformed to 100.) Afterward, some type of ordinary least squares regression is performed on the data, often at the respondent level, to identify the utilities, or coefficients, of the individual elements. The regression analysis does not “recognize” the differences in meaning between the persuasion model and the interest model. The additive constant of the interest model, again an estimated parameter, can be interpreted as the conditional probability of a respondent being interested in the concept, *i.e.* giving a rating of 7–9, if there are no elements present. As before, this additive constant is a purely estimated parameter, but it constitutes a good baseline. The individual utilities, $k_1.k_{36}$, from the regression model show the conditional additive probabilities of the individual elements. A high positive utility value in the interest model means that more respondents will find the concept interesting when the element is inserted into the concept. A high negative utility in the interest model means that more respondents will find the concept disinteresting when the element is inserted into the concept.

Researchers interested in the mind of the individual consumer, *e.g.,* experimental psychologists, typically pay more attention to the equation when the dependent variable is the rating itself and thus base their analyses on the persuasion model. These researchers trace their intellectual orientation to experimental psychology. In contrast, researchers who trace their intellectual heritage to sociology are more interested in whether or not a particular respondent is a member of a class, such as the class of concept acceptors. These researchers, especially market researchers, focus more on the interest model because that model deals with membership in a subgroup.

HOW RDE PROMOTES ANALYSIS OF INTERACTIONS USING "SCENARIOS"

When product developers and marketers talk about the requirements for their products, they often begin by saying that they are interested in what particular features "work with" or "go with" other features that have already been selected (Krieger, Green and Wind, 2004). For example, the developer might decide to incorporate certain ingredients with known health benefits, or to create the product under a specified brand name. The research issue now becomes the discovery of what other elements "work" within that scenario. In other cases, where the decisions have not even yet been made, it's important to learn what elements will work with each of a number of alternative brand names. Each option that the developer wants to make as a given for the product constitutes its own scenario. Thus, each brand name would be treated as a separate scenario.

The RDE-based conjoint analysis can deal with these scenarios and thus reveal the nature and magnitude of implicit interactions. When the objective is to find what other product variables *go with* a product feature, the researcher can answer this problem by considering only those concepts that have this product feature and then building a model using those concepts in order to estimate the contribution of the other product elements. That is, the researcher simply divides the full set of concepts from all the respondents into the relevant strata. Each option of the ingredient to be forced in will, in turn, define its own stratum. For each stratum, the researcher can then build a model showing how the other elements contribute. Of course, the set of concepts chosen is a subset of the full set of concepts and thus some of the balance among the appearance of concept elements, so carefully designed for the full set of concepts, will be skewed. Generally, however, there are sufficient number of distinct concepts generated by the approach so that this resulting imbalance does not become too much of a problem.

This analysis by "scenario" works in a conjoint study because of the nature of the task. In the conjoint study, the elements are mixed and matched so each element behaves as a free agent. The background against which the element appears also varies. Thus from one concept to another the respondent is being exposed to different subset of elements. There is no obvious mental set being developed so the respondent might be said to react instinctively to new combinations at an almost intuitive level.

Within this framework of permuted designs, it becomes fairly straightforward to combine the records of all of the respondents into one very large data file. The approach to identifying these scenarios follows four steps:

Create a resource data file: comprising the concepts for each respondent, followed by the rating. Each record (line) of this very large data file comprises the respondent identification number. For a study say of six categories and six elements per category, the file comprises a further set of 36 columns, one column reserved for each of the 36 elements. For a specific respondent and for a specific concept, the row shows the number "1" if the particular element appears in that concept and the number "0" if the particular element is missing from that concept. Clearly most of the numbers in the row will be "0" because the concept does not contain most of the elements. Finally, the row contains the rating assigned by the respondent (1–9 scale), as well as the transformed rating (1–6 transformed to 0; 7–9 transformed to 100). Table **1** shows an example of what this partial matrix might look like for the first three experimentally designed concepts evaluated by three respondents.

Table 1: Example of Raw Data for the First Three Experimentally Designed Concepts Evaluated by Three Respondents (Partial Data).

Resp.	Con.#	A_1	A_2	F_3	F_4	F_5	F_6	Rate	Int.
001	1	0	0	1	0	0	0	2	0
001	2	0	0	0	0	0	0	5	0
001	3	0	0	0	0	0	0	2	0
002	1	0	0	0	0	0	0	3	0
002	2	1	0	0	0	0	1	8	100
002	3	0	0	1	0	0	0	5	0
233	1	0	0	0	1	0	0	8	100
233	2	0	1	0	0	0	0	4	0
233	3	0	0	0	0	0	1	8	100

Select one independent variable (element): For example, one can choose any of the elements in the category A and then select only those specific cases where this particular element was present. We see this in Table **1** by looking at the appropriate column corresponding to the element and finding the row, *i.e.* the case, where there is a "1". By repeating the selection process for every element in the particular category, the researcher splits the whole data set for all the respondents into several different layers. For *N* elements in a category (*e.g.,* category A), there will be ($N + 1$) layers—one per each element in the category and one for cases when the category is absent, because the experimental design is set up so that some of the concepts are missing elements from the particular category.

Run separate regressions, one for each level in a variable (*i.e.* a separate regression for each element in the category): Working with all of the elements except those in the category used to create the stratum, run a separate regression model relating the presence/absence of the elements to the interest value. In the example of category A, we would regress interest *versus* the presence/absence of $B_1 \ldots F_6$ for each element in category A. There would be seven such regressions, one regression for each element in category A (*i.e.* A_1, $A_2 \ldots A_6$) and one regression for those cases where no element from category A actually appeared in a test concept. The regression would thus comprise 30 independent variables, not 36. There would be seven such regression equations, one for each value of category A plus one regression when no elements from category A are present at all. In effect, this analysis would show how each of the element in category A "drives" the utilities of the remaining elements. If category A were to be key ingredients, then the researcher would uncover how every element performs if the concept is identified with a specific key ingredient name.

A complete analysis of the scenarios generates separate regression analyses for each element in each category. For C equally sized categories and M elements per category, this complete analysis would generate $C \times (M + 1)$ regression models. Thus for six categories with six elements in each one, there would be 6×7 or 42 regression models. By looking at each set of regression models, one per category, the researcher would be able to identify the implicit interaction effects of each element with every other element in a different category.

DISCOVERING PAIRWISE SYNERGISMS AND SUPPRESSIONS THROUGH EXPERIMENT

There are great benefits to being able to address the issue of synergism and suppression for concept development and messaging. Like products and concepts, messages comprise components. Unlike products, these components are by their very nature "discrete", "on-off", "absent or present". There is almost no way that someone will know what "synergizes" together, basically because no researcher has ever tested that many combinations to measure this synergism. Thus, any method that discovers the existence of synergisms among these hundreds of combinations adds to the ability of research to uncover new opportunities in messaging.

Our approach to discovering and quantifying pairwise synergisms is fairly straightforward and based on rigorous statistical considerations. We follow the four steps listed below.

Create the full matrix of concept data: *Concept Elements* (columns) × *Concept* (rows). This is the same starting matrix that we used to work with scenarios (see Table **1**).

Expand the matrix to include all pairwise interactions as additional columns: By the nature of the experimental design, only elements from different categories can appear together in the same concept. Therefore for C categories, each with E elements and with all categories having the same number of elements, there are ((C)

$\times (C - 1)/2)$ distinct pairs of categories and $E \times E$ pairs of elements for each distinct pair of categories. The product of the category pairs and the element pairs defines the number of all possible synergisms to be discovered. This number mounts up very rapidly. For example, with four categories and with six elements per category there are six pairs of categories (*i.e.* $4 \times 3/2$), each pair of categories comprising 36 pairs of elements. The total is 216 combinations of elements that might synergize. When we increase the categories to a larger design comprising six categories, each with six elements, we increase the number of pairs from 216 to 540.

First, create the linear model and afterward determine which synergistic pairs increase predictability beyond the simple linear terms: The first approach forces in the linear terms and subsequently allows additional terms, *i.e.* synergistic pairs, to enter the already-developed equation when those terms contribute significant additional predictability to the equation. The criteria for entering the linear equation and adding additional terms can be made stringent or lax using the mechanics of stepwise multiple linear regression (SPSS, 2004). With a stringent criterion applied to the regression modeling, very few additional predictor terms will enter the equation. In contrast, applying a lax criterion lets many additional terms enter the equation. One of the issues is the degree to which one can be lax, allowing many additional terms to enter in order to capture the interactions, yet be true to the original model.

Second, explore all possible interactions: even before creating the model, simply to see which candidate terms show synergy. This second strategy explores the combinations, but does not yet allow in any predictors. This strategy identifies those combinations that *could* add additional predictability without, however, making any decision about which terms to add. This exploration is also easily done with stepwise multiple linear regression, first by forcing in all the linear terms and second by setting the criteria so stringently for entering the equation that no synergistic combinations can ever enter. The result is a sense of the potential strength of each of the synergistic terms to enter and augment the linear equation. The result of this strategy is a ranking of all the terms by strength of interaction (given by the F ratio of the term) and the direction (given by the partial correlation, with a positive correlation meaning synergism and a negative correlation meaning suppression).

EXPLORING WHAT CAN BE LEARNED THROUGH IMPLICIT AND EXPLICIT INTERACTIONS

The best way to understand what can be learned is by means of a set of experiments. This chapter presents the results of two experiments, originally designed to understand the drivers of acceptance from the point of view of product features using conjoint analysis. However, the data sets also provide the opportunity to explore interactions.

The first and smaller data set is about the consumer response to different concepts dealing with a variety of sweeteners, including sugar, high fructose corn syrup and four high potency sweeteners (*e.g.,* aspartame). This first data set was developed in conjunction with work on consumer response to different sweetener types, to determine whether the sweetener name affects acceptance. Other variables in this study were primarily emotional in nature, so this first data set is a good introduction to interactions among variables where one category (sweetener name) may carry significant emotional baggage. The second and very much larger data set deals with consumer responses to the features of a cookie product. This second data set was designed almost entirely as a product development exercise. All of the elements are product features with relatively little emotion attached to the elements, other than as a very relevant tag line to a specific feature.

EMPIRICAL STUDY—MESSAGING FOR SWEETENERS

Background

This study was originally commissioned by a sweetener manufacturer in order to understand how consumers would respond to different types of messaging about sweeteners. The sweetener category is highly competitive, with both caloric and high potency sweeteners vying for the consumer's wallet (Hardin and Marquardt, 1967). Caloric sweeteners include the gold-standard sucrose and the less expensive, very popular sweetener high fructose corn syrup (HFCS), which is substituted for sucrose in order to reduce costs. Often the taste quality deteriorates slightly, but formulators modify the composition of the product to ameliorate some of the loss. High potency sweeteners appeal to a different part of the market: those who want low calories and are willing to give up some of the taste. There is intense competition among these high potency sweeteners, based upon the need to recoup development and patent costs. New sweeteners are being sought all the time because of the desire to identify a sweetener with the taste profile of sucrose, the "gold standard", *i.e.* with no side-tastes. Most high potency sweeteners have off-tastes that may or may not be successfully masked by product formulations.

Most consumer respondents are not particularly interested in sweeteners *per se*. They buy products, with the sweetener being part of the product. It's more meaningful to test concepts about products with the sweetener being part of the concept. Thus, the messaging study was constructed in terms of an orange beverage, with the sweetener elements appearing in the concepts along with other elements more relevant to the beverage and to emotional rewards.

CONCEPT ELEMENTS, EXPERIMENTAL DESIGN, FIELD EXECUTION

The consumer respondents were recruited by e-mail invitation. Those opting to participate evaluated 40 different concepts. The experimental design was permuted so that each respondent evaluated a different set of 40 combinations. The concept elements appear in Table **2**, along with the utility values. These values are the conditional probabilities that a respondent will be interested in the orange beverage concept. The data show average utilities from 38 individual-level respondent models. As noted above, although the study had originally been commissioned to understand the mind of the consumer respondent with respect to different sweeteners, it was important to embed these sweeteners in a context in which the sweeteners themselves would not attract undue attention. The relative utility values of the sweeteners would still emerge because they were parts of the concept. Their utilities could be compared across natural and high potency sweeteners.

Table 2: Categories, Elements and Utility Values for the Sweetener Study, Positioned as a Study of Orange Beverages.

Elements		**Utility**
Additive constant		43
Category A: Introduction/Positioning		
A_1	Introducing a fabulous new orange beverage.	3
A_2	A brand new orange beverage.	–3
A_3	A new beverage that cools your mouth.	–3

Table 2: cont….

A_4	An orange beverage that reminds you of Florida.	2
A_5	Introducing an outstanding new product—oranges!	–2
A_6	Ignite your senses with a burst of orange!	2
Category B: Taste Promise		
B_1	Drink it up—you'll love the taste!	0
B_2	A cool sensation every time you sip.	4
B_3	It warms your mouth—and your heart.	–12
B_4	It goes down smooth.	–3
B_5	Icy cold—a full sensation!	4
B_6	Brings you back to the islands.	0
Category C: Sweetener		
C_1	Sugar to give it a nice sweet taste	–24
C_2	Saccharin to give it a nice sweet taste	–44
C_3	Sucralose to give it a nice sweet taste	–3
C_4	Acesulfame K to give it a nice sweet taste	–43
C_5	Aspartame to give it a nice sweet taste	–40
C_6	High sweet corn syrup to give it a nice sweet taste	–31
Category D: Use		
D_1	Drink it for yourself	–3
D_2	Buy it for your whole family	5
D_3	A great taste for your kids	7
D_4	Try it soon with your friends	2
D_5	Try it soon with your husband or wife	–3
D_6	Great taste for the whole family	4

THE UTILITIES OF ELEMENTS AND THE IMPLICIT SYNERGISMS FROM SCENARIOS

The utility is a measure of performance. High utilities mean that the element does well, *i.e.* the element drives consumer interest. Low elements, near zero and negative, mean that the element does poorly, *i.e.* the element reduces consumer interest. The utilities in Table **2** suggest that the different sweeteners perform poorly except for sucralose. Furthermore, most of the remaining concept elements perform neither well nor poorly. The only exception is element B_3 (*It warms your mouth—and your heart*) with a utility of –12. Table **3** suggests that there is very little differentiation among the elements when we look at the entire database of concepts.

Table 3: Seven Scenarios, Created from Concepts with the Six Different Sweeteners and the Seventh Scenario Created from Concepts Missing a Sweetener Statement (Category C_0).

Element code	Element text	Linear	Category C						
			C_0	C_1	C_2	C_3	C_4	C_5	C_6
Additive constant		43	48	6	–1	27	2	5	17
Category A: Positioning									
A_1	Introducing a fabulous new orange.	3	12	5	6	9	9	–7	–17
A_2	A brand new orange beverage.	–3	19	–2	0	–14	–2	2	–11
A_3	A new beverage that cools your mouth.	–3	1	8	12	2	0	–7	–1
A_4	An orange beverage that reminds you…	2	17	–7	7	–1	1	1	–3
A_5	Introducing … new product—oranges!	–2	3	7	8	23	–2	–7	–15
A_6	Ignite your senses with a burst of orange!	2	8	12	–1	12	–1	–6	0
Category B: Taste Promise									
B_1	Drink it up—you'll love the taste!	0	–1	11	–5	13	1	3	–12
B_2	A cool sensation every time you sip.	4	–10	21	–4	19	5	9	–9
B_3	It warms your mouth—and your heart.	–12	–24	0	–3	–9	–1	0	–8
B_4	It goes down smooth.	–3	–12	–3	–1	–13	–2	0	18
B_5	Icy cold—a full sensation!	4	3	11	2	10	3	7	–1
B_6	Brings you back to the islands.	0	2	–14	–4	4	–1	8	–3
Category D: Use									
D_1	Drink it for yourself	–3	–12	–3	6	30	4	1	–8
D_2	Buy it for your whole family	5	–3	21	4	20	1	–7	3
D_3	A great taste for your kids	7	0	22	11	15	–1	–4	0
D_4	Try it soon with your friends	2	–14	–5	10	–4	–1	–4	6
D_5	Try it soon with your husband or wife	–3	–19	8	–2	30	3	1	14
D_6	Great taste for the whole family	4	–6	0	–1	1	–1	–6	5

Let us go further, however. How do these different elements perform when we look at concepts stratified by sweeteners? Can we find an interaction between sweetener type and the specific message? Let us follow the decomposition approach for scenarios discussed above. We stratify the raw data along the six different levels for sweetener (elements C_1–C_6) to create six different data sets for analysis, as well as a seventh data set comprising those concepts that contained no mention of sweetener. Recall that the experimental design comprises concepts that are absent all mention of a category of elements. This strategy of zero conditions ensures that we can run meaningful regression models and that the utilities have absolute value. For each category, in turn, that strategy creates a stratum of concepts with the category absent.

Four results emerge from the analysis by scenario:

1. The additive constant varies by scenario. The additive constant measures basic interest in the concept without any elements being present. In a scenario analysis, the additive constant measures the interest in the element on which the stratum is founded. We start with C_0, the stratus where no sweetener is mentioned. Concepts for the stratum absent a sweetener show a relatively high additive constant (48). From this we infer that the basic interest in no sweetener is high (48). Now add the different sweeteners. The basic interest in the stratum for sucralose is 27. We infer from that that sucralose is acceptable as well. Further down, we find high fructose corn syrup with a basic acceptance of 17 achieved by its stratum. For the remaining high potency sweeteners, Table **2** tells us that their strata generate much lower, even negative utilities. Thus the interest in these elements is very low.

2. We're really interested in how well the same element performs in different strata. When the same element has a high utility in one stratum and a low utility in another stratum, we have evidence for interaction. The sweetener defining the stratum must be interacting with that particular element.

3. Following this logic, we now discover which particular elements go well with each sweetener.

4. With the scenarios, it becomes straightforward to create optimal concepts for each of the sweeteners. Let us contrast three optimal concepts—one for sugar; the second for high fructose corn syrup, also a caloric sweetener that is less expensive than sugar; and the third for saccharin. The tonality of the sugar concept is slightly different from the tonality of the high fructose corn syrup (HFCS) and both are modestly different from the tonality of saccharin.

a. For sugar, the more natural and familiar caloric sweetener, the optimal concept is:

   Sugar to give it a nice sweet taste

   Ignite your senses with a burst of orange!

   A cool sensation every time you sip.

   Buy it for your whole family or A great taste for your kids

b. For the less familiar caloric sweetener, HFCS, the optimal concept is:

   High sweet corn syrup to give it a nice sweet taste

   Ignite your senses with a burst of orange!

   It goes down smooth or Try it soon with your husband or wife

c. For saccharin, the well-known high potency sweetener, the optimal concept is…

   Saccharin to give it a nice sweet taste

   A new beverage that cools your mouth

   Icy cold—a full sensation!

   A great taste for your kids or Try it soon with your friends

ASSESSING SYNERGISMS DIRECTLY FROM UTILITY MODELS COMPRISING LINEAR TERMS AND INTERACTIONS

The direct assessment of synergism for the sweeteners can be done by stepwise regression analysis, which begins by forcing in the linear terms and then adding pairwise interaction terms to increase the goodness of fit of the model to the data. Depending upon the criteria used to add the interaction terms, the resulting equation can incorporate just a few interaction terms, or all the terms. The number of interaction terms to include in the equation is a function of the researcher's criteria.

We assess the impact of adding few interaction terms *versus* adding all interaction terms in a relatively simple fashion, following these steps to create three models:

Model 1: Create a pure linear model with no interactions, using the ordinary least squares regression model.

Model 2: Create a model that forces in all of the linear terms and afterward allows *some* of the interactions to enter the model. Operationally, this means forcing in linear terms, but later allowing interactions to enter if the criterion to enter is high. We look at two levels of stringency. The more stringent is an *F* ratio of 4.0 to enter. The less stringent is an *F* ratio of 2.0 to enter. The former excludes more terms, the latter excludes fewer terms.

Model 3: Create a model that allows *all* of the interactions to enter. Operationally, this means making the *F* ratio 0 for a term to enter, after the linear terms have been forced in.

We judge how reasonable the interaction terms are by comparing the coefficients of the 24 elements in the simple linear model (no interactions—Model 1 above) with the coefficients for the same 24 linear terms in the two models with interactions (Model 2 or more stringent and Model 3 or not stringent at all).

One of the continuing issues in looking at interactions emerges when we realize that adding the effects of the interactions will actually change the linear coefficients or utilities. We certainly don't want to have a shifting landscape of utilities, just because we add interaction terms. Rather, we'd like the interaction terms to add more predictability, but at the same time leave relatively untouched the utilities we previously computed from the linear model.

Let us now look at this "perturbing" effect of interactions, *i.e.* how the addition of interaction terms manages to upset the values of the linear terms that we assume to be the case and which we computed before we even dealt with interactions. We see the scatter plot of the coefficients in Fig. **2**. The abscissa shows the utility values of the 24 elements when the models were first created from the individual respondent data and then averaged across all 38 respondents. These are our original data points. The ordinate shows the utility values for the *linear portion of the interaction model*, for the three criteria: *Fewest* number of interactions or most stringent ($F = 4.0$); *Moderate* ($F = 2.0$); and *Most* or least stringent ($F = 0.0$). We see clearly that as more interaction terms enter the equation, the utility values for the 24 concept elements in the linear portion of the model (without interactions) depart from what we observed initially. Although we want to identify as many significant interactions as possible, it is best to limit the number of interactions so that the linear portion of the model with interactions is as similar as possible to the linear portion of the model without interactions. In this way, adding interactions does not change our interpretation of how the elements perform.

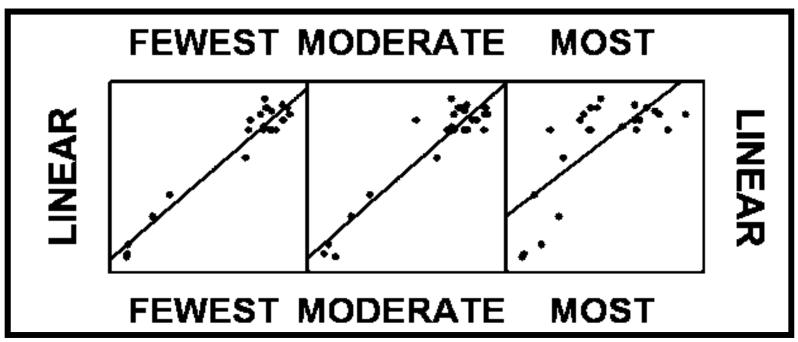


Figure 2: Scatterplot matrix of the coefficients in the orange beverage study. Three panes correspond to the number of interactions terms.

A better understanding of the nature of the synergisms appears in Table **4**, which shows the outcome of stepwise regression. The criteria were the most stringent (F to enter = 4.0; F to remove = 3.9). All linear terms were forced in, whether or not they were significant. The interaction terms were afterward allowed to enter the equation, but only if they added significant predictability to the model. The results in Table **4** suggest that utility values around 6 or so begin to become very significant, as the probability of their being really 0 diminishes to below 0.10 (given by the P statistic). By this operationally defined method to select significant interactions, we end up with a limited number of highly significant interactions, which in turn exhibit large coefficients (utility values beyond +15 or –15).

Table 4: Specifics of the Equation Relating the Interest and Interactions in the Orange Beverage Study.

Term	**Coefficient (utility)**	***P* value**
Additive constant: 42		
Linear terms in model (all forced in)		
A_1	1	0.70
A_2	1	0.69
A_3	3	0.45
A_4	**6**	**0.06**
A_5	–1	0.83
A_6	6	0.06
B_1	5	0.18
B_2	0	0.97
B_3	**–6**	**0.08**
B_4	–5	0.14
B_5	**6**	**0.08**
B_6	4	0.19
C_1	**–27**	**0.00**
C_2	**–39**	**0.00**
C_3	–5	0.23
C_4	**–39**	**0.00**
C_5	**–39**	**0.00**
C_6	**–32**	**0.00**
D_1	1	0.78
D_2	4	0.18
D_3	–1	0.88
D_4	–2	0.58
D_5	–1	0.79
D_6	0	0.94
Pairwise interaction terms (all very highly significant)		
A_1*D_3	**25**	**0.01**
A_2*C_3	**–22**	**0.02**
A_2*D_3	**27**	**0.00**
A_4*B_2	**–21**	**0.03**
E_5*D_3	**29**	**0.00**
A_6*B_1	**–31**	**0.01**
B_2*C_1	**23**	**0.02**
B_2*C_5	**19**	**0.04**
B_4*C_6	**31**	**0.00**

Table 4: cont…..

B_6*D_1	**–30**	**0.01**
C_3*D_1	**25**	**0.01**
C_3*D_5	**30**	**0.00**

Note. Significant utilities are in boldface; only significant interactions are shown.

By itself, the results in Table **4** simply show which terms interact with each other to generate positive synergisms and which terms interact to generate negative synergisms (suppression). We finish this analysis of interactions by looking at different pairs of elements that interact with each other to determine how a linear model might have treated the pairwise combination *versus* how a model with interactions might have treated the same pairwise combination. The results should not be surprising given the large interaction values in Table **4**. We see the differences between the interaction model and the pure linear model in Table **5**, which compares the results using the interaction model to the results using the pure linear model from the original individual-level analysis (Table **2**). For a number of these combinations, knowledge of strong pairwise interactions could dramatically increase or decrease the total interest score for a pair of elements in a test concept. The effect could be 15–20 points. Knowing which pair of elements interact may, in some cases, make a big difference in how one selects the elements for a concept.

Table 5: Performance of Combinations of Elements in the Sweetener Study that Interact Synergistically.

Cross term	**Elements**	**Text**	**Interaction model**			**Pure linear model**		**(Interaction model) – (linear model)**
			Linear term	**Interaction**	**Sum + 42[a]**	**Linear term**	**Sum + 43[b]**	
Pairs of orange beverage elements that appear to act synergistically								
B_4*C_6	B_4	It goes down smooth…	–5	31	36	–3	9	27
	C_6	High sweet corn syrup to give it a nice sweet taste	–32			–31		
C_3*D_5	C_3	Sucralose to give it a nice sweet taste	–5	30	66	–3	37	29
	D_5	Try it soon with your husband or wife	–1			–3		
A_5*D_3	A_5	Introducing an outstanding new product—oranges!	–1	29	69	–2	48	21
	D_3	A great taste for your kids	–1			7		

Table 5: cont….

A_2*D_3	A_2	A brand new orange beverage.	1	27	69	–3	47	22
	D_3	A great taste for your kids	–1			7		
A_1*D_3	A_1	Introducing a fabulous new orange beverage.	1	25	67	3	53	14
	D_3	A great taste for your kids	–1			7		
C_3*D_1	C_3	Sucralose to give it a nice sweet taste	–5	25	63	0	40	23
	D_1	Drink it for yourself	1			–3		
B_2*C_1	B_2	A cool sensation every time you sip.	0	23	38	4	23	15
	C_1	Sugar to give it a nice sweet taste	–27			–24		
B_2*C_5	B_2	A cool sensation every time you sip.	0	19	22	4	7	15
	C_5	Aspartame to give it a nice sweet taste	–39			–40		
Pairs of orange beverage elements that appear to act suppressively								
A_4*B_2	A_4	An orange beverage that reminds you of Florida.	6	–21	27	2	49	–22
	B_2	A cool sensation every time you sip.	0			4		
A_2*C_3	A_2	A brand new orange beverage.	1	–22	16	–3	40	–24
	C_3	Sucralose to give it a nice sweet taste	–5			0		
B_6*D_1	B_6	Brings you back to the islands.	4	–30	17	0	40	–23
	D_1	Drink it for yourself	1			–3		
A_6*B_1	A_6	Ignite your senses with a burst of orange!	6	–31	22	2	45	–23
	B_1	Drink it up—you'll love the taste!	5			0		

Notes: [a]Additive constant in the interaction model; [b]additive constant in the pure linear model. The table shows the pair of elements, the linear coefficients, the interaction coefficient, the sum and comparison to the pure linear model, as well as the difference.

It is important to determine how these synergisms come about. Sometimes they come from elements with modest negative utilities in the linear portion, but very strong positive synergisms. *The ideal combination would be some positive utilities for each element in the linear portion of the model and dramatic positive synergism in the interactions*. The pair of elements coming closest to that ideal combination is $A_1 + D_3$:

"Introducing a fabulous new orange beverage". + *"A great taste for your kids".*

DISCUSSION

The discussion of the scenario and interaction approach follows two distinct paths. The first path considers the process of developing messages and products. The second path considers statistical issues.

PROCESS CONSIDERATIONS

Using the scenario approach (implicit interactions) for product development and for messaging. A key goal in product development is to identify the combination of features that "go together". Traditionally, this identification emerges from questionnaires. Respondents most frequently scale interest. Far less frequently, respondents rate "going together" for pairs of features. There's a reason for this. For N features that could combine into 2-tuples there are N ratings for one feature and $(N)(N - 1)/2$ ratings required for all pairs. The list of combinations can become very long and not particularly easy to analyze, although the question about combinations of features continues to be raised in meetings on product development. Combinations of features are very relevant for developers.

We might imagine using the same stepwise approach for messaging rather than product development. For messaging, the objective is to convince consumers about one's message, rather than identify the specific products. Typically, messaging research is done through focus groups rather than by stepwise and systematic analysis of the alternatives. This difference in approach between product development and messaging development comes from the recognition that product development requires the selection of appropriate physical variables. In product development, it is important to select the right stimuli. In contrast, in messaging development the creation of messages may be done with the same belief that it is a scientific process, but messages are in truth created more out of one's intuition and opinion, rather than being created through science. Nonetheless, it is quite possible to automate the scenario approach, identify core sets of messages and then through the scenario analysis and regression identify what particular additional messages complement the basic core message.

Using the explicit interactions for product development and message development. *Explicit interactions* more typically represent what the statistician calls interactions. Developers of both products and messages always look for combinations that provide “more” just by virtue of appearing together. This chapter shows that significant explicit interactions can add as many as 10–30 points to interest, or can subtract a similar amount if the combination is suppressive. It is not necessary to work with synergistic combinations, but rather the synergism is an extra benefit that should be taken advantage of if possible. In contrast, it is important to avoid the suppressive combinations. Knowledge that a specific combination is likely to lose more points than the components contribute helps the developers to craft better products and messages.

One of the ongoing issues in the study of interactions is the “*why”*. One might ask why certain combinations do unusually well or unusually poorly. Is this simply a statistical aberration? Are the interactions repeatable? Do they make sense and is there a pattern underlying these interactions that can teach us about how we respond to the components of products and messaging? As this is the first study of its type that appears to have worked with interactions, those questions cannot be answered as yet.

STATISTICAL CONSIDERATIONS

Limits of the method—number of interactions to be sampled. One of the key aspects of the approach here is the ability to combine the data from many respondents into one very long data file. For example, with each respondent evaluating 48 combinations and with 500 respondents the data set comprises 24,000 records. A benefit of this large data set is the ability to test many elements and many interactions. There is no concern about the degrees of freedom. With six categories, each comprising six elements, there are 36 main effects to be estimated and an additional 15 × 36 or 540 interaction effects. With eight categories, each comprising eight elements, there are 64 main effects to be estimated along with 28 × 64 or 1792 interaction effects. Furthermore, the interaction effects enter the equation only when they can be shown to add significant predictive power to the variables already in the equation. Thus, the main technical problems confronting the approach are the number of predictors available in today’s regression packages. The number of cases is almost never a problem. Happily, the ever-increasing power of off-the-shelf, shrink-wrapped regression modules makes the computation limits something that inevitably will be overcome, usually within a fairly short time after the need for more computational power is recognized for a workhorse statistical program, such as regression modeling.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This chapter is based on the previously published Gofman, A. (2006). Emergent scenarios, synergies and suppressions uncovered within conjoint analysis. *Journal of Sensory Studies, 21*(4): 373–414; and Gofman A. (2009). *Experimentation-based product development in mature food categories: Advancing conjoint analysis approach*. Tartu, Estonia: Tartu University Press.

REFERENCES

Carmone, F., Green, P. and Jain, A. (1978). Robustness of conjoint analysis: some Monte Carlo results. *Journal of Marketing Research, 15*(2), 300-303.

Green, P. and Srinivasan, V. (1990). Conjoint analysis in marketing: new developments with implications for research and practice. *The Journal of Marketing, 54*(4), 3-19.

Hardin, D. and Marquardt, R. (1967). Increasing Precision of Market Testing. *Journal of Marketing Research, 4*(4), 396-399.

Krieger, A., Green, P. and Wind, Y. (2004). *Adventures in Conjoint Analysis: A Practitioner's Guide to Trade-off Modeling and Applications.* Retrieved from http://marketing.wharton.upenn.edu/ideas/pdf/Green/Monograph/Adventures-Introductory%20Materials.pdf

Moskowitz, H., Gofman, A. and Beckley, J. (2006). Using high-level consumer-research methods to create a tool-driven guidebook and database for product development and marketing. *Journal of Sensory Studies, 21*(1), 54-100.

Moskowitz, H., Porretta, S. and Silcher, M. (2005). *Concept research in food product design and development*: Blackwell Pub.

SPSS® 13.0 (2004). Brief guide. Chicago, IL: SPSS Inc.

Wittink, D. and Cattin, P. (1989). Commercial use of conjoint analysis: An update. *The Journal of Marketing, 53*(3), 91-96.

Wittink, D., Vriens, M. and Burhenne, W. (1994). Commercial use of conjoint analysis in Europe: results and critical reflections. *International Journal of Research in Marketing, 11*(1), 41-52.

CHAPTER 4

Putting RDE on the R&D Map: A Survey of Approaches to Consumer-Driven New Product Development

Alex Gofman*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: This chapter summarizes different consumer-driven approaches to new product development with their advantages and disadvantages. The analysis identifies opportunities for development in the field.

Keywords: New product development, consumer-driven innovation, rule developing experimentation, conjoint analysis.

OVERVIEW OF CONSUMER APPROACHES TO NPD

In today's business environment, a continuous supply of novel products is essential to retaining competitive advantage. Jaeger, Rossiter, Wismer and Harker (2003), Buisson (1995) and Costa and Jongen (2006) state that new product development (NPD) is often recommended as a suitable strategy to build competitive advantage and long-term financial success of companies in today's global food markets by approaching the development of new products in a more structured manner. Lord (1999), Meulenberg and Viaene (1998), Trail and Grunert (1997) and van Trijp and Steenkamp (1998) demonstrate that product innovation helps maintain growth, spread market risk, enhance the company's stock market value and increase competitiveness.

There are many methods to create new products. They are not all equal. (Cooper, 1979), Burchill and Fine (1997), Ford and Sternman (1998) and Stewart-Knox and Mitchell (2003) research progressively more sophisticated approaches to consumer-driven NPD, which emerged in past decades, although a majority of them concentrate on industrial development rather than mature and saturated markets such as food products.

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

Innovation takes on many faces. The first is the nature of the innovation, reactive or proactive, respectively. Buisson (1995) reported that many food corporations utilize a *reactive approach* to NPD, in which companies try to market what was developed through marketing efforts rather than developing what consumers wanted in the first place. For the food industry the reactive approach should not surprise. Food technology cannot easily modify the base products to create dramatically new innovations, at least in the ordinary course of events. On the other hand, argue Urban and Hauser (1993), continuous collection and aggregation of the consumer-related data, along with taking consumer needs as the starting point of product innovation efforts, create the opportunity for a *proactive approach* to NPD, even in the absence of radically new innovations.

Innovation also varies by its source. Traditionally, *i.e.* in the years before 1980, innovation came from the top, from the laboratory of the corporation marketing the product. Competition changed a lot of that. *User-driven innovation* was first described by von Hippel (1976, 1978, 1988), who documented a number of cases where customers modified or adapted existing products according to their own needs before the industry did. (Grunert *et al.* 2008) argue that today the term has an extended meaning and is applicable to all forms of consumer involvement in the innovation process. A number of related concepts exist in the literature, including *early customer integration* (Gassmann and Wecht, 2005), *participatory design* (Mayhew, 1999) and *user-centered development* (Ketola and Ahonen, 2005). Grunert *et al.* (2008) further suggest a variation of the term—*user-oriented innovation*, which is defined as a process towards the development of a new product or service in which an integrated analysis and understanding of the users' wants, needs and preference formation play a key role. Costa and Jongen (2006) and Grunert and Valli (2001) demonstrate that users can be both direct customers and end users.

Consumer-led NPD was introduced in Urban and Hauser (1993), Urban *et al.* (1997) and further developed in van Trijp and Steenkamp (1998) and Vriens, Loosschilder, Rosbergen and Wittink (1998) along with others. Consumer-led NPD is one market-oriented innovation process that takes into account the current and future needs of consumers, along with understanding what really adds value in products that satisfy those needs.

Consumer-led developments are not random. Rather, they follow the market. It's not clear whether this is by design or simply by evolution. Urban and Hauser (1993) suggest that the process does have a structure with these steps (see Fig. **1**):

1. need identification,

2. idea development to address the need,

3. product development to substantiate the idea and

4. product's market introduction to communicate the fulfillment of a need.

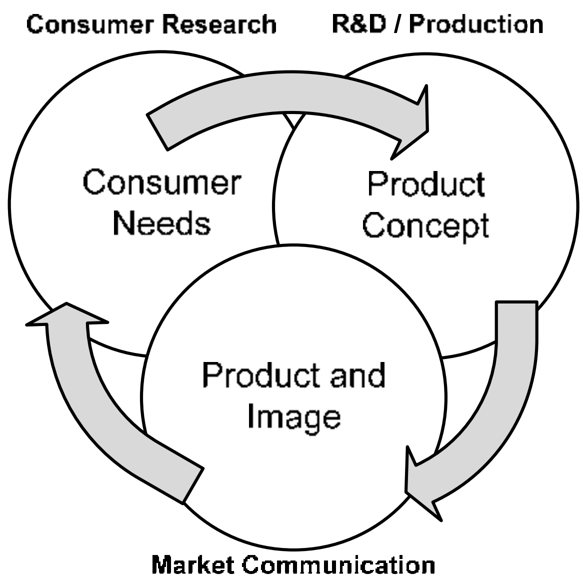


Figure 1: The consumer-led NPD (adapted from Urban and Hauser, 1993).

Whether the process emerges from trial and error in an evolutionary way or design top down is not necessarily clear from today's literature. Observation of companies will reveal that for the most part the process evolves over time. Yet the key point is that there is a process. The effective ability to translate subjective needs (*e.g.,* healthy or convenient) into objective product specifications is essential for the realization of the satisfaction of consumer wishes through the development of a new product.

It's not enough that knowledge becomes products. That's only half of the issue. The other half is how to sell the products and services. Concurrently, marketing strategies must be initiated to communicate the existence of a new or improved product, which satisfies consumer needs in a distinctive and superior way. Dahan and Hauser (2002), Grunert, Baadsgaard, Larsen and Madsen (1996), Urban and Hauser (1993), van Trijp and Steenkamp (1998) and Wind and Mahajan (1997) all demonstrate in an evidence-based way that such a consumer-led approach to product development often greatly increases the likelihood of success of innovation processes.

The often evolutionary rather that conscious design of consumer-led NPD can be a weakness, especially in slower-moving industries. For instance, Fuller (1994), van Trijp and Steenkamp (1998) and Stewart-Knox and Mitchell (2003) all point out the somewhat sequential rather than concurrent, overlapping, or iterative nature of consumer-led NPD in the food industry. This structure, working as it does through years of history, becomes in effect a weakness and a considerable obstacle to success.

There are "fixes" however. Wheelright and Clark (1992) propose the *funnel* approach, an improvement to consumer-led NPD. Cusumano and Selby (1995) address the shortcoming though the *spiral* approach. Both these approaches start with a broad range of ideas from several sources that are later winnowed to a few high-potential concepts, some of which will, in turn, be ultimately developed and launched.

Lord (1999), Moskowitz, Porretta and Silcher (2005) and Beckley, Foley, Topp, Huang and Prinyawiwatkul (2007) suggest further steps in consumer-driven approaches. Their industrial orientation, rather than pure academic orientation in the world of business science make their suggestions less than rigorous. They do not provide theoretical foundations, perhaps because the latter two authors come from the world of solution providers and business practice, rather than from the world of academic business science.

CONCEPTS—THE RAW MATERIAL OF CONSUMER-LED NPD

A critical step in new product development selects from multiple and potential product concepts those specific ones that the firm would consider introducing to the market. At this practical step of selection two different ways of doing things have sprung up. These are practices, each valid in its own right, but radically different from each other.

The two main approaches to the concept selection and choosing a design prior to detailed engineering and prototyping are, respectively, the *method of controlled convergence* introduced in Pugh (1996) and *concept engineering* described in Burchill (1992). The foregoing assumes one concept. There is a third way, not always attractive to business practitioners who want closure or who want the "one solution". Srinivasan, Lovejoy and Beach (1997) argue that, depending on the cost of developing each concept into a customer-ready prototype, it would be optimal to carry *multiple product concepts* into the prototyping and testing phase and only later, with more knowledge, select the best of those designs. Fig. **2** shows how they think of the process, where multiple options are carried along until the final decision is made, presumably with more knowledge at the later stage.

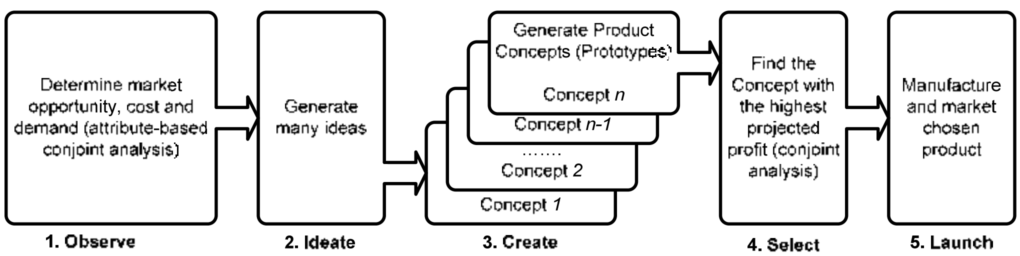


Figure 2: NPD process consisting of five stages. Stages 1, 3 and 4 could utilize conjoint analysis. Source:(Dahan and Srinivasan, 2000) (author's rendering).

At a practical level, when we deal with concepts and prototypes, the question is how many? It's one thing to work with one concept and one prototype. On the other hand, when there are many alternatives that can be chosen, the company must support many centers of effort. What is the payout for each? And, more important, how can the problem be addressed with today's knowledge of business strategies? Dahan and Mendelson (2001) recommend the *extreme value theory* to determine the optimal number *n* of parallel prototypes and compare the approach to the more common and perhaps less productive, iterative strategy beloved by practitioners who want progress to proceed in a linear fashion. Dahan and Mendelson argue that it would be worthwhile to increase the optimal number of prototypes in those situations where time to market is short and the cost of prototyping is low. There is a great deal of agreement with their suggestions. Stevens and Burley (1997), van Kleef, van Trijp and Luning (2005) and Wansink (2005) also point out the need to work with many prototypes to develop a successful product.

Despite the cost, moving ahead with multiple options through the process has other benefits as well, besides a greater chance of arriving at the best solution. Markets change. The complexity of today's markets, the interlinked nature of companies, consumers and choices make it necessary to have alternatives "in the hopper" at all times. And so, even 15 and 20 years ago there was a sense that the parallel path was best. This notion of parallel paths was elaborated in 1995 by Iansiti, who underlined its two big advantages: added flexibility to respond to market and technological shifts, along with shortening the total product development time in case of such shifts.

EVEN BEFORE CONCEPTS—WHAT DOES THE CUSTOMER WANT?

Getting to know customer needs is becoming increasingly popular as the pre-work for concepts. It should not come as a surprise that there are continuing, competitive approaches to this discipline. Academics study the discipline but it is left to the business practitioner and solution provider to make such knowledge come alive.

We begin with the *Voice of the Customer* (VOC), analyzed in Adams (2004). VOC is a proactive and continually innovative process, or better, a set of disparate processes, which capture the requirements/feedback from the customer (internal or external). The goal is to provide the customers with the best in class service/product quality by soliciting their feedback and suggestions. Carson, Gilmore, Perry and Gronhaug (2001) define the goal of VOC as one that aims to detect the stated and unstated needs or requirements of the customer. These unstated needs and requirements can be captured in a variety of ways: direct discussion or interviews, surveys, focus groups, customer specifications, observation, warranty data, field reports, complaint logs. A whole business of solution providers has emerged who specialize in the techniques falling loosely under the rubric of VOC.

The range of VOC is great, but occasionally quite limited. Zaltman and Coulter (1995) demonstrate how VOC identifies the specific quality attributes needed for a supplied component or material to incorporate in the process or product. Yet VOC isn't prophetic, nor is it creative. According to Leonard (2002), VOC is limited to selection from the options known to consumers. When customers are asked to make new product recommendations or elaborate on topics they have limited or no knowledge of, they tend to run into at least two kinds of blocks. The first is the human tendency to fixate on the way products or services are normally used, making people unable to imagine alternative functions. The second is that people may not be able to conceive of a solution because they might have contradictory needs.

In previous years, most of the attention in consumer research was focused on assessing products and concepts that were "complete", "polished" and ready to go. More recently, attention has shifted from this late stage of development to a much earlier stage, where the options for development are open ("fuzzy front end"; Reid and de Brentani, 2004). Yet Dahan and Mendelson (2001) suggest that the "fuzzy innovation stage" is often ignored by consumer researchers in favor of the more defined, later stages of development where there are metrics, standard procedures and recognized analytical approaches. Process and exploitation take over from exploration, perhaps because the former is more natural in business, the latter more risky and as such less natural. As a result, the important first stage of product development is poorly executed, leaving the product not optimized to survive the competition (Toubia, 2004).

The idea generation and testing methods that belong naturally to the fuzzy front end of NPD provide information to designers about the customer utility, technical feasibility and cost of a new product, respectively. Allocating resources at this early stage poses a challenge, considering the inherent uncertainty. In particular, uncertainty arises from imperfect information about customers and markets, undiscovered or untested product designs and technologies and the challenges in executing and delivering an ideal design. Concept tests help resolve this uncertainty and encompass most methods used to measure the performance of new products and processes along those dimensions affecting profitability. Methods such as *VOC*, *contextual inquiry* (Raven and Flanders, 1996), *lead user analysis* (von Hippel, 2005), *conjoint analysis*, *Kano methods* (Shen, Tan and Xie, 2000) and *Pugh concept selection* (Dahan and Hauser, 2002) clarify and focus the fuzzy front end of NPD. For example, Raven and Flanders (1996) show that contextual inquiry and VOC methods help identify key product attributes. Green and Srinivasan (1990a, 1990b) point out that conjoint analysis quantifies the importance of each attribute and the degree of price sensitivity, providing designers with the information needed to optimize attribute-level trade-offs for specific market segments.

VIABLE STRATEGIES

With all of the interest in systematic approaches to NPD, the key interest is operationally viable strategies. Just how have people done it? What is required? What are successes and failures in this area of early stage reconnaissance, design, development and testing?

This area, which got the attention of researchers such as von Hippel (2005), provides us with a range of methods through which lead users or less professionalized communities can take on a role as designers to interact with product developers in companies. Research described in Grunert *et al.* (2008) show that lead users are those individuals whose present needs will become widespread in a market months or years ahead. Since lead users are intuitively familiar or better "in touch" with conditions that will become the near future for most others, these lead users can become the need-forecasting laboratory for market research.

We begin with the latest stage, research with customer-ready prototypes. Testing these prototypes reveals technical problems and opportunities that might have been overlooked in the design's conceptual phase (Srinivasan *et al.* 1997). When the effort has been made to create the prototypes, other methods can be used as well. *Product archeology* (Ulrich and Pearson, 1998) and *design for manufacturability* (Boothroyd, Dewhurst and Knight, 1994) identify potential cost savings through product simplification and process improvement, although they have limited applicability.

By the time the prototype is created, a lot of money has been spent. The research is typically diagnostic. Any design is typically redesign. Let's go back one step, before the effort has been invested to create the prototype. Now we are at the concept stage, or the manual/computer design step. At this step, one process that is increasingly becoming accepted is *systematic recombination* of RandD-meaningful features. This mix/match strategy can, in turn, be enhanced by including the consumer preferences (Rabino *et al.* 2007; Varian, 2003). The latter implies that the combinations are directed, rather than being random (Varian, Farrell and Shapiro, 2004). Moskowitz *et al.* (2005) argue that this *combinatorial innovation* is especially applicable to food product NPD and could jump-start the innovation process even in mature categories.

QUALITY FUNCTION DEPLOYMENT—FORMALLY INVOLVING CONSUMERS IN REQUIREMENTS

Models of the processes for the food industry should emphasize involving the consumer from the start of the process and should integrate technology and marketing efforts (Stewart-Knox and Mitchell, 2003). One such model is the *House of Quality,* the first of four phases within *Quality Function Deployment* (QFD; Costa, Dekker and Jongen, 2001). Originally developed for improving the automotive industry in Japan (Akao, 2004), QFD was applied more than a decade ago by van Trijp and Steenkamp (1998).

The QFD is a planning process for the design of new products (Akao, 2004; Costa *et al.* 2001; Garcia *et al.* 2007). QFD provides a systematic method to translate "customer requirements" into design and process parameters. These latter parameters are the so-called company requirements (King, 1992). In doing so, QFD helps companies to reduce two types of risks (Holmen and Kristensen, 1998). The first risk is the failure of the product specification to correspond to the needs and wants of a predetermined target group of customers. The second risk is the failure of the final product to correspond to the original product specifications.

The QFD is based on the belief that products should be designed to reflect customers' demands and needs. Therefore, the project requires top management commitment and organizational support. Beyond that support there must be an inherent clarity of the structure. The project objectives must be specified. Finally, the development is not in the domain of one group alone, later to be passed to another group. At the start all groups must participate. This requires creating a cross-functional team. That team should comprise members from all the company's functional areas involved in NPD, in market introduction and consumer product testing, respectively (Costa *et al.* 2001).

Finally, it is worth noting that QFD provides a statistically based tool to advance NPD, but it may encounter some problems in the real world of product development. Such problems are especially common when QFD is applied to raw materials that vary. *Interactions* between attributes can play a decisive role in achieving consumer satisfaction. Garcia *et al.* (2007) argue that raw materials have a natural predisposition for variation that does not fit well with the somewhat inflexible character of QFD charts described in Dekker and Linnemann (1998). Many food ingredients show interactions and some HOWs could affect more than some WHATs. These interactions and the large list of customer demands are often seen as the major bottlenecks of using QFD in new food product development (ibid).

OTHER APPROACHES TO INVOLVING THE CONSUMER IN NPD

Despite the increasing array of tools offered by academics, by solution providers and developed at the company through trial and error, there is great room for improvement. One of the major and most obvious gaps in consumer-driven food product development is the lack of clear and concrete guidelines for effective implementation in everyday company practices (Nijssen and Frambach, 2000). The methods are known, at least in theory and taught by practitioners to those who would implement the methods in practice. Yet it's often not clear just what one should do! This deficiency in "marching orders" to deploy the methods is felt mostly at the early phases of the development process—opportunity identification and opportunity definition—which are at once the less structured yet important determinants for the success of new products (Nijssen and Lieshout, 1995).

There are other methods, therefore, which find their place when companies grapple with what to do in NPD. It's not just the traditional large-scale approaches that are adopted. A variety of other methods, old and new, find application. For example anderson-Connell, Ulrich and Brannon (2002) describe companies trying to use traditional qualitative methods to improve the process such as *focus groups.* Prescott, Young, O'Neill, Yau and Stevens (2002) report the use of *Food Choice*

Questionnaire (FCQ), which assess the relative importance of factors thought to be important motives in food choice: Health, Mood, Convenience and Sensory Appeal. These factors can be modeled by conjoint analysis as well (Deliza, 2003).

Statistics isn't limited to modeling. As computer technology and capability spread wide to practitioners, advanced methods that were once in the domain of experts now find themselves widely used. In the 1960s, psychometricians began to explore different ways to represent stimuli on a geometrical map such that stimuli lying close together geometrically were somehow related to each because they were similar to each other. These mapping approaches took a while to become accepted. Over time, however, the geometric representation of stimuli became part and parcel of the research world. Mapping found its way into the world of NPD as well.

In recent years, a class method, which has gained popularity, is the now well-accepted *preference mapping techniques*. Whereas preference mapping is used to address a wide range of problems, it is particularly suited to aid new product development (Greenhoff and MacFie, 1994). Mapping helps show the opportunities in the market, literally. On the other hand, representing opportunities is not the same thing as identifying exactly what to do. Although mapping combines combine sensory profiles and acceptability data to identify the sensory attributes driving consumer acceptance, there are no specific next steps of an operationally defined nature. Hence the applications of mapping are limited, beyond its use in representation. Murray and Delahunty (2000) deal with mapping and the issue in application, focusing on the specific case of cheddar cheese.

VALUE OF THE CONSUMER INPUT

A poll of author researchers dealing with consumer needs will come up with the observation that understanding consumer needs is important. The list of such researchers gets bigger every year and has for decades. Just a small sampling of papers will show this. Cooper and Kleinschmidt (1994), Grunert *et al.* (1996), Urban and Hauser (1993) all write forcefully that understanding consumer needs and reacting effectively to them is one of the most important correlates of product development success.

Yet there remain questions regarding the value of consumer focus in NPD. It has been stated that consumer-driven development activities, by following closely consumer needs, encourage incremental innovation in detriment to the development of truly new products. Atuahene-Gima (1995), Ortt and Schoormans (1993), van Trijp and Schifferstein (1995) and Wind and Mahajan (1997) all point to the main argument sustaining this view, namely that consumers cannot be expected to provide needs about products or technologies that are yet unknown to them.

To answer those approaches, other methods have been developed. These can only be mentioned here because each requires a treatise by itself. These methods include *consumer-idealized design* (Ciccantelli and Magidson, 1993), *problem* and *lead-users design* (Ortt and Schoormans, 1993; von Hippel, 1986), *beta-testing* (Kaulio, 1998) and *information-acceleration.* The latter is interesting because it places consumers in future technological scenarios (Urban and Hauser, 1993), simulating reality as closely as possible, in an affordable way.

AT THE END OF THE DAY, WHERE ARE WE IN TERMS OF METHODS?

Unfortunately, the business demands of new product development all too often hinder the accumulation of knowledge, much less wisdom. We don't necessarily become smarter, despite the success or failure that we encounter. This failure to accumulate knowledge, in a systematic manner, about new product launches (both successful and failed) exerts a negative effect on the effectiveness of NPD. Yet, again and again, researchers studying the new product process point to how much more can be achieved by simply systematizing knowledge. The list of papers is impressive: Deshpandé and Zaltman (1982), Inkpen and Dinur (1998), Li and Calatone (1998), Davenport and Klahr (1998), Arnett, Menon and Wilcox (2000) and De Luca and Atuahene-Gima (2007) all reiterate, again and again, that integrating experience, knowledge and wisdom through knowledge management in a systematic, continuous way from RandD all the way to the launch could substantially improve the chances for success. Sherman, Berkowitz and Souder (2005) demonstrate that the combined effect of RandD and marketing integration and knowledge management in the form of recording, retrieving and reviewing information from past projects amplify corporate capabilities, ranging from proficiency in developing prototypes, to proficiency in launching. As an extra bonus, such systematic knowledge adds to the company's core competency and makes the reworking of prototypes far easier and more productive.

DISCUSSION AND CONCLUSIONS

There are several emerging trends in consumer-driven experimental design-based approaches to NPD:

1. Most of the research in the area of NPD concentrates on relative young industrial processes and products. Yet and in contrast, there are many fields (*e.g.,* food industry) where the majority of the product development are forced to happen in mature categories. The traditional limitations of the resources allocated for NPD in mature product categories present special academic and practical interests.

2. Some businesses and societies have a misperception that there is not much place for innovation in mature product categories and the efforts should be concentrated only on NPD in emerging areas. This chapter demonstrates that the notion is incorrect and there is always a place for innovation, even in mature categories that are especially difficult for the developers.

3. There is an unfulfilled need for low-cost, effective and reliable testing of new product features on the concept level, which should be systematic and disciplined.

4. Corporations frequently rely on the instincts of designers to develop new products. The approaches described in this chapter further develop the ideas introduced by a large school of researchers who believe in the power of knowledge of consumer preferences in product development as well as in consumer-driven systematic and disciplined approached to NPD.

CONFLICT OF INTEREST

None declared.

ACKNOWLDEGEMENTS

This chapter is partially based on Gofman A. (2009). Experimentation-based product development in mature food categories: Advancing conjoint analysis approach. Tartu, Estonia: Tartu University Press.

REFERENCES

Adams, J. (2004). BofA's 'Customer Voice' sounds off in ePayments. *Direct Access, 17*(5), 37.

Akao, Y. (2004). *Quality function deployment: Integrating customer requirements into product design*. New York, NY: Productivity Press.

Anderson-Connell, L., Ulrich, P. and Brannon, E. (2002). A consumer-driven model for mass customization in the apparel market. *Journal of Fashion Marketing and Management, 6*(3), 240–258.

Arnett, D., Menon, A. and Wilcox, J. (2000). Using competitive intelligence: Antecedents and consequences. *Competitive Intelligence Review, 11*(3), 16–27.

Atuahene-Gima, K. (1995). An exploratory analysis of the impact of market orientation on new product performance: A contingency approach. *Journal of Product Innovation Management, 12,* 275–293.

Beckley, J., Foley, M., Topp, E., Huang, J. and Prinyawiwatkul, W. (2007). *Accelerating new food product design and development* (Institute of Food Technologists Series). Hoboken, NJ: Wiley-Blackwell.

Boothroyd, G., Dewhurst, P. and Knight, W. (1994). *Product design for manufacturability and assembly*. New York, NY: Marcel Dekker.

Buisson, D. (1995). Developing new products for the consumer. In D. Marshall (Ed.), *Food choice and the consumer* (pp. 182–215). Cambridge, MA: Chapman and Hall.

Burchill, G. (1992). *Concept engineering: The key to operationally defining your customers' requirements.* Cambridge, MA: Center for Quality Management.

Burchill, G. and Fine, C. (1997). Time *versus* market orientation in product concept development: Empirically based theory generation. *Management Science, 43,* 465–478.

Carson, D., Gilmore, A., Perry, C. and Gronhaug, K. (2001). *Qualitative marketing research.* London, UK: SAGE Publications.

Ciccantelli, S. and Magidson, J. (1993). From experience: Consumer idealized design—involving consumers in the product development process. *Journal of Product Innovation Management, 10,* 341–347.

Cooper, R. and Kleinschmidt, E. (1996). Winning businesses in product development: The critical success factors. *Research Technology Management, 39,* 18–29.

Cooper, R. (1979). The dimensions of industrial new product success and failure. *The Journal of Marketing, 43*(3), 93-103.

Costa, A. and Jongen, W. (2006). New insights into consumer-led food product development. *Trends in Food Science and Technology, 17,* 457–465.

Costa, A., Dekker, M. and Jorgen, W. (2001). Quality function deployment in the food industry: A review. *Trends in Food Science and Technology, 11,* 306–314.

Cusumano, M. and Selby, R. (1995). *Microsoft secrets*. New York, NY: The Free Press.

Dahan, E. and Hauser, J. (2002). The virtual customer. *Journal of Product Innovation Management, 19*(5), 332-353.

Dahan, E. and Mendelson, H. (2001). An extreme-value model of concept testing. *Management Science* (Design and Development Issue), *47*(1), 102–116.

Dahan, E. and Srinivasan, V. (2000). The predictive power of internet-based product concept testing using visual depiction and animation. *Journal of Product Innovation Management, 17*(2), 99-109.

Davenport, T. and Klahr, P. (1998). Managing customer support knowledge. *California Management Review, 40*(3), 195–208.

Dekker, M. and Linnemann, A. (1998). Product development in the food industry. In W. Jongen and M. Meulenberg (Eds.), *Innovation of food production systems: Product quality and consumer acceptance* (pp. 67–86). Wageningen, the Netherlands: Wageningen Press.

Deliza, R. (2003). Use of computer-generated images and conjoint analysis to investigate sensory expectations. *Journal of Sensory Studies, 18,* 465–486.

De Luca, L. and Atuahene-Gima. K. (2007). Market knowledge dimensions and cross-functional collaboration: Examining the different routes to product innovation performance. *Journal of Marketing, 71,* 95–112.

Deshpandé, R. and Zaltman, G. (1982). Factors affecting the use of market research information: A path analysis. *Journal of Marketing Research, 19*(1), 14–31.

Ford, D. and Sternman, J. (1998). Dynamic modeling of product development processes. *Systems Dynamics Review, 14,* 31–68.

Fuller, G. (1994). *New food product development.* From concept to marketplace. Boca Raton, FL: CRC Press.

Garcia, M., de Carlos, P., de Felipe, I., Briz, J., Morais, F. and Navarro, M. (2007, February 15–17). Quality function deployment: Can improve innovation efficiency in the food industry? In IEFISD (Ed.), *Proceedings of The 1st International European Forum on Innovation and System Dynamics in Food Networks*. Innsbruck-Igls, Austria: IEFISD.

Gassmann, O. and Wecht, C. (2005). Early customer integration into the innovation process. In EIASM (Ed.), *Proceedings of the 12th International Product Development Conference.* Copenhagen, Denmark: EIASM.

Grunert, K., Jensen, B., Sonne, A., Bruns, K., Byrne, D., Clausen, C., *et al.* (2008). User-oriented innovation in the food sector: relevant streams of research and an agenda for future work. *Trends in Food Science and Technology, 19*(11), 590-602.

Gold, A., Malhotra, A. and Segars, A. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185–214.

Green, P. E. and Srinivasan, V. (1990a). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5,* 103–123.

Green, P. E. and Srinivasan, S. (1990b). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing, 54*(4), 3–19.

Greenhoff, K. and MacFie, H. (1994). Preference mapping in practice. In MacFie and Thomson (Eds.), *Measurement of food preferences* (pp. 137–166). London, UK: Blackie Academic and Professional.

Grunert, K. and Valli, C. (2001). Designer-made meat and dairy products: Consumer-led product development. *Livestock Production Science, 72,* 83–98.

Grunert, K., Baadsgaard, A., Larsen, H. and Madsen, T. (1996). *Market orientation in food and agriculture*. Boston, MA: Kluwer.

Grunert, K. G., Jensen, B. B., Sonne, A.-M., Brunsø, K., Byrne, D. V., Clausen, C.,. Scholderer, J. (2008). User-oriented innovation in the food sector: Relevant streams of research and an agenda for future work. *Trends in Food Science and Technology, 19*(11), 590–602.

Holmen, E. and Kristensen, P. (1998). Supplier roles in product development: Interaction *versus* task partitioning. *European Journal of Purchasing and Supply Management, 4*(2–3), 185–193.

von Hippel, E. (1976). The dominant role of users in the scientific instrument innovation process. *Research Policy, 5,* 213–239.

von Hippel, E. (1978). Successful industrial products from customer ideas. *Journal of Marketing, 40*(1), 39–49.

von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science, 32*(7), 791–805.

von Hippel, E. (1988). *The sources of innovation*. New York, NY: Oxford University Press.

von Hippel, E. (2005). *Democratizing innovation*. Cambridge, MA: MIT Press.

Iansiti, M. (1995). Shooting the rapids: Managing product development in turbulent environments. *California Management Review, 38,* 37–58.

Inkpen, A. and Dinur, A. (1998). Knowledge management processes and international ventures. *Organizational Science, 9*(4), 454–468.

Jaeger, S., Rossiter, K., Wismer, W. and Harker, F. (2003). Consumer-driven product development in the kiwifruit industry. *Food Quality and Preference, 14,* 187–198.

Kaulio, M. (1998). Customer, consumer and user involvement in product development: A framework and review of selected methods. *Total Quality Management, 9,* 141–149.

Ketola, P. and Ahonen, M. (2005). Learning and innovating between users and developers within new technology development. In EIASM (Ed.), *Proceedings of 12th International Product Development Conference*. Copenhagen, Denmark: EIASM.

King, J. (1992). Corporate leadership in manufacturing. *Planning Review, 20*(5), 19–20.

van Kleef, E., van Trijp, H. and Luning, P. (2005). Consumer research in the early stages of new product development: A critical review of methods and techniques. *Quality and Preference, 16*(3), 181–201.

Leonard, D. (2002). The limitations of listening. *Harvard Business Review, 80*(1), 93.

Li, T. and Calatone, R. (1998). The impact of market knowledge competence on new product advantage: Conceptualization and empirical examination. *Journal of Marketing, 62*(4), 13–29.

Lord, J. (1999). New product failure and success. In A. Brody and J. Lord (Eds.), *Developing new food products for a changing marketplace* (pp. 55–86). Lancaster, PA: Technomic Publishing Company.

Mayhew, D. (1999). *The usability engineering lifecycle*. San Francisco, CA: Kaufman.

Meulenberg, M. and Viaene, J. (1998). Changing food marketing systems in western countries. In W. Jongen and M. Meulenberg (Eds.), *Innovation of food production systems: Product quality and consumer acceptance* (pp. 5–36). Wageningen, the Netherlands: Wageningen Press.

Moskowitz, H. R., Porretta, S. and Silcher, M. (2005). *Concept research in food product design and development*. Ames, IA: Blackwell Professional Publishing.

Murray, J. and Delahunty, C. (2000). Mapping consumer preference for the sensory and packaging attributes of cheddar cheese. *Food Quality and Preference, 11*(5), 419–435.

Nijssen, E. and Frambach, R. (2000). Determinants of the adoption of new product development tools by industrial firms. *Industrial Marketing Management, 29,* 121.

Nijssen, E. and Lieshout, K. (1995). Awareness, use and effectiveness of models and methods for new product development. *European Journal of Marketing, 29,* 27–44.

Ortt, R. and Schoormans, J. (1993). Consumer research in the development process of a major innovation. *Journal of the Market Research Society, 35,* 375–387.

Prescott, J., Young, O., O'Neill, L., Yau, N. and Stevens, R. (2002). Motives for food choice: A comparison of consumers from Japan, Taiwan, Malaysia and New Zealand. *Food Quality and Preference, 13,* 489–495.

Pugh, S. (1996). Concept selection: A method that works. In D. Clausing and R. Andrade (Eds.), *Creating innovative products using total design: The living legacy of stuart pugh* (pp. 167–176). Upper Saddle River, NJ: Addison-Wesley.

Rabino, S., Moskowitz, H., Katz, R., Maier, A., Paulus, K., Aarts, P., Beckley, J. and Ashman, H. (2007). Creating databases from cross-national comparisons of food mind-sets. *Journal of Sensory Studies, 22*(5), 550–586.

Raven, E. and Flanders, A. (1996). Using contextual inquiry to learn about your audiences. *ACM SIGDOC Asterisk Journal of Computer Documentation, 20*(1), 1–13.

Reid, S. and de Brentani, U. (2004). The fuzzy front end of new product development for discontinuous innovations: A theoretical model. *Journal of Product Innovation Management, 21*(3), 170–184.

Saguy, I. and Moskowitz, H. (1999). Integrating the consumer into new product development. *Food Technology, 53*(8), 68–73.

Shen, X., Tan, K. and Xie, M. (2000). An integrated approach to innovative product development using Kano's model and QFD. *European Journal of Innovation Management, 3*(2), 91–99.

Sherman, D., Berkowitz, D. and Souder, W. (2005). New product development performance and the interaction of cross-functional integration and knowledge management. *Journal of Product Innovation Management, 22*(5), 399–411.

Stevens, G. and Burley, J. (1997). 3,000 Raw Ideas = 1 Commercial Success! *Research-Technology Management*, May–June, 16.

Stewart-Knox, B. andMitchell, P. (2003). What separates the winners from the losers in new food product development? *Trends in Food Science and Technology, 14,* 58–64.

Srinivasan, V., Lovejoy, W. and Beach, D. (1997). Integrated product design for marketability and manufacturing. *Journal of Marketing Research, 34,* 154–163.

Toubia, O. (2004). *New approaches to idea generation and consumer input in the product development process*. Boston, MA: MIT Press.

Trail, B. and Grunert, K. (1997*). Product and process innovation in the food industry.* London, UK: Chapman and Hall.

van Trijp, H. and Schifferstein, H. (1995). Sensory analysis in marketing practice: Comparison and integration. *Journal of Sensory Studies, 10,* 127–147.

van Trijp, H. and Steenkamp, J. (1998). Consumer-oriented new product development: Principles and practice. In W. Jongen and M. Meulenberg (Eds.), *Innovation of food production systems: Product quality and consumer acceptance* (pp. 37–66). Wageningen, the Netherlands: Wageningen Pers.

Trott, P. (2008). *Innovation management and new product development*. Upper Saddle River, NJ: Prentice Hall.

Ulrich, K. and Pearson, S. (1998). Assessing the importance of design through product archaeology. *Management Science, 44*(3), 352–369.

Urban, G. and Hauser, J. (1993). *Design and marketing of new products* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Urban, G., Hauser, J., Qualls, W., Weinberg, B., Bohlmann, J. and Chicos, R. (1997). Information acceleration: Validation and lessons from the field. *Journal of Marketing Research, 34,* 143–153.

Varian, H. (2003). Innovation, components and complements. UC Berkeley. Retrieved July 31, 2008, from http://www.almaden.ibm.com/coevolution/pdf/varian_paper.pdf

Varian, H., Farrell, J. and Shapiro, C. (2004). *The economics of information technology: An introduction*. Cambridge, MA: Cambridge University Press.

Vriens, M., Loosschilder, G., Rosbergen, E. and Wittink, D. (1998). Verbal *versus* realistic pictorial representations for including design-attributes in conjoint analysis. *Journal of Product Innovation Management, 15,* 455–467.

Wansink, B. (2005). Consumer profiling and the new product development toolbox: A commentary on van Kleef, van Trijp and Luning. *Food Quality and Preference, 16,* 217–221.

Wheelright, S. and Clark, K. (1992). *Revolutionizing product development*. New York, NY: The Free Press.

Wind, J. and Mahajan, V. (1997). Issues and opportunities in new product development: An introduction to a special issue. *Journal of Marketing Research, 34,* 1–12.

Zaltman, G. and Coulter, R. (1995). Seeing the voice of the customer: Metaphor-based advertising research. *Journal of Advertising Research, 35*(4), 35–51.

CHAPTER 5

RDE in Concept Research: An Empirical Demonstration

Alex Gofman[1,*] and Tõnis Mets[2]

[1]Moskowitz Jacobs Inc., White Plains, New York, USA and [2]Centre for Entrepreneurship, University of Tartu, Tartu, Estonia

Abstract: The chapter illustrates a practical RDE implementation for market research applications demonstrated with IdeaMap®.NET (http://www.ideamap.net) online RDE tool. Using a case study, the chapter follows steps of the RDE process with explanations of data and findings.

Keywords: Rule developing experimentatioN (RDE), conjoint analysis, interactions.

INTRODUCTION

This chapter demonstrates the rule developing experimentation (RDE) process, step by step, utilizing IdeaMap®.NET online Software-as-a-Service (SaaS) application. The process is illustrated on the examples used in Gofman (2009) and Gofman and Moskowitz (2009).

RDE encompasses six basic steps, which we will illustrate in this chapter:

1. Prepare raw materials
2. Create test stimuli (concepts)
3. Collect data from consumers
4. Analyze the data
5. Identify pattern-based latent segments and interactions
6. Apply the emergent rules to create new products, services, offerings

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

Step 1. Prepare Raw Materials

RDE operates with *elements*—units of decomposition such as features, messages, or other items of experimentation, combined into concepts. Elements are concise descriptive sentences or other media creating individual elements of the products grouped into attributes of similar features. Examples of the elements are: "*Comes in a sealed metal can*", "*Low-fat*" and so forth. Elements in RDE become the so-called levels of the variables, in the language of experimental design.

Attributes (also called variables, silos or categories)—a group of related units of decomposition—common elements from which the concepts (conceptual prototypes) are built. Examples of attributes are: "Flavors" (with possible elements such as "Orange-flavored", "Chocolate-flavored", "Strawberry-flavored"); "Fat Content" (with elements such as "Fat-free", "Low-fat"). In RDE, a maximum of one element from each attribute can be selected to appear in a test concept. There are many test concepts with no elements from an attribute.

The RDE process begins with the basic architecture of the test concept, *i.e.* the number of attributes and the number of elements per attribute. Fig. **1** shows a screen shot wherein the user can specify the specific number of attributes and the number of attributes per element. The figure shows the screen shot for the RDE tool, known as IdeaMap®.NET. For this example, we choose a design comprising six attributes/six elements in each attribute. Therefore, there are 36 elements in the RDE study, with each a separate element.

The basic foundation of RDE lies in the elements. These elements have "meaning". That is, the elements are independent ideas, which could function as single phrases that an ordinary person can easily understand. Elements come from different sources: brainstorming, focus groups, competitive analysis, *etc.* Table **1** shows the elements. At the right side of Table **1**, we see "utility" values, which will be obtained by experiment. We will deal with the utilities in the sections below. For right now, it suffices simply to recognize the different elements and get a sense of how "user-friendly" these elements turn out to be.

The elements then are entered into RDE program such as IdeaMap®.NET (Fig. **2**). The process is very similar to working with a spreadsheet application such as Excel.

Table 1: The Additive Constant and the Utilities for the Total Panel (Gofman and Moskowitz, 2009).

	Elements	**Total**
	Base size	*439*
	Additive constant	*34*
Attribute A – What does it look like and feel like (appearance, texture)?		
A1	Soft and chewy…just like homemade	8
A3	Soft and chunky…for an extra special treat	8
A4	Oversized chunks of dark chocolate to sink your teeth into	7
A6	Jumbo size…for when you want a little extra	5
A5	Bite size for a quick indulgence	4
A2	Crisp and crunchy…perfect for dunking	1
Attribute B – What does it contain (ingredients)?		
B1	Real creamery butter for a rich, indulgent taste	7
B5	Made with only the freshest ingredients…eggs, milk, butter	3
B6	Sweetened with natural fructose for a healthy indulgence	2
B3	For a healthy source of protein….made with unpasteurized egg whites	1
B4	Made with unprocessed whole grain flour…keeping all the goodness in	1
B2	Made with canola oil which helps lower blood cholesterol levels	-1
Attribute C – What 'healthful features' does it offer?		
C4	0 grams trans fat and cholesterol free…a heart friendly cookie	5
C5	A high-fiber cookie that boosts your energy level and leaves you feeling full	5
C6	With no trans fats or preservatives…for a healthy snack you can feel good about giving your kids	5
C1	Calcium enriched for strong bones	4
C2	Low carb…when you're counting carbs and looking for a great snack	2
C3	With added iron and isoflavones… a cookie that not only tastes good but is good for you	2
Attribute D – What flavors does it feature?		
D3	Dark Belgian chocolate, Swiss milk chocolate and bittersweet chocolate…simply irresistible	11
D6	Oatmeal…for old fashioned goodness and packed with nutrition	6
D4	Rich and creamy peanut butter…for those who love an old favorite	4
D5	Vanilla flavored…a traditional favorite	-2
D1	Comes in spicy flavors…cinnamon, nutmeg and allspice	-3
D2	Comes in cool, citrus flavors…orange, lemon and lime, perfect for a ladies afternoon tea	-9
Attribute E – How is it packaged or stored?		
E3	Comes in resealable bags.take out only what you want	6
E6	Comes in a crush-proof box…no more broken cookies	6

Table 1: cont….

E1	All cookies are individually wrapped for freshness	4
E4	Available in variety packs…three of your favorite varieties in one box	4
E2	Comes with a stay fresh inside wrapper…just twist and seal	2
E5	Available in decorative tins…the perfect gift idea	2
Attribute F – Where in the store is it found and how is it merchandised?		
F1	In the bakery section of your supermarket…always fresh	2
F2	In the frozen foods section of your supermarket…just thaw and serve	1
F3	In the frozen foods section of the supermarket…bake and serve…hot from the oven	5
F4	In the refrigerated section of your supermarket…just thaw and serve	1
F5	In the refrigerated section of your supermarket…just bake and serve	3
F6	Sold in bulk…enough to please a crowd	2

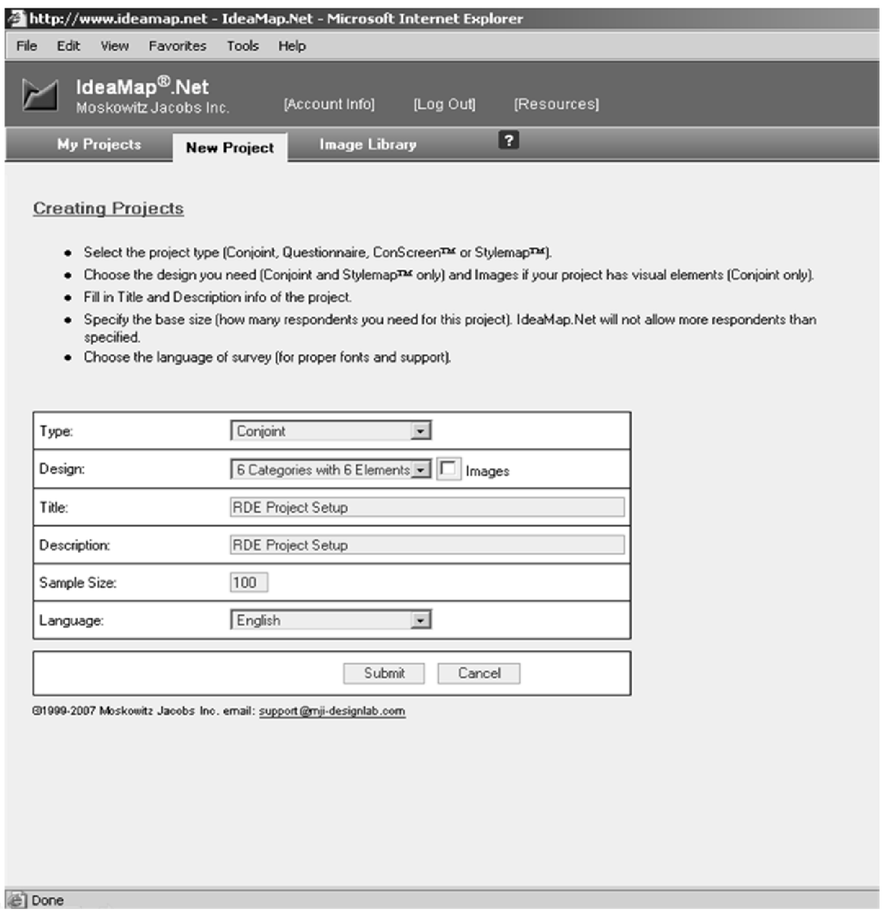


Figure 1: Screen capture of the project creation step in IdeaMap®.NET. Drop-down menu shows available designs.

Step 2. Create Test Stimuli (Concepts)

Our example comprises six attributes with six elements in each. We select an appropriate original experimental design (Table **2**). The experimental design ensures that each element appears an equal number of times and appears in a statistically uncorrelated way with any other element. Each individual will evaluate combinations from a unique experimental design. However, each unique experimental design is really an isomorphic permutation of the base design. The permutation module operates on the original or base experimental design. Chapter 2 shows two individual permuted designs, one for respondent 1, the second for respondent 2. The two designs are isomorphic to each other.

Table 2: Original (Source) Experimental Design Utilized in the Project for Individual Permutations.

Unit	Var^1	Var^2	Var^3	Var^4	Var^5	Var^6
1	x_0^1	x_1^2	x_1^3	x_0^4	x_4^5	x_1^6
2	x_1^1	x_6^2	x_0^3	x_0^4	x_0^5	x_4^6
3	x_0^1	x_2^2	x_2^3	x_4^4	x_5^5	x_0^6
4	x_5^1	x_2^2	x_0^3	x_3^4	x_5^5	x_0^6
5	x_3^1	x_0^2	x_3^3	x_4^4	x_0^5	x_1^6
6	x_3^1	x_1^2	x_0^3	x_2^4	x_4^5	x_0^6
7	x_4^1	x_0^2	x_1^3	x_5^4	x_2^5	x_0^6
8	x_2^1	x_0^2	x_3^3	x_0^4	x_2^5	x_0^6
9	x_0^1	x_6^2	x_1^3	x_0^4	x_4^5	x_0^6
10	x_0^1	x_1^2	x_4^3	x_0^4	x_3^5	x_2^6
11	x_0^1	x_5^2	x_1^3	x_4^4	x_6^5	x_0^6
12	x_0^1	x_3^2	x_4^3	x_0^4	x_0^5	x_5^6
13	x_3^1	x_0^2	x_6^3	x_6^4	x_0^5	x_5^6
14	x_2^1	x_4^2	x_0^3	x_1^4	x_0^5	x_2^6
15	x_0^1	x_2^2	x_1^3	x_2^4	x_0^5	x_4^6
16	x_3^1	x_5^2	x_0^3	x_6^4	x_5^5	x_0^6
17	x_2^1	x_0^2	x_6^3	x_0^4	x_0^5	x_5^6
18	x_0^1	x_4^2	x_6^3	x_5^4	x_6^5	x_0^6
19	x_5^1	x_4^2	x_0^3	x_0^4	x_1^5	x_3^6
20	x_5^1	x_0^2	x_5^3	x_6^4	x_0^5	x_3^6
21	x_4^1	x_0^2	x_5^3	x_0^4	x_3^5	x_5^6
22	x_2^1	x_3^2	x_0^3	x_6^4	x_0^5	x_6^6
23	x_1^1	x_0^2	x_2^3	x_5^4	x_6^5	x_0^6
24	x_6^1	x_0^2	x_4^3	x_0^4	x_5^5	x_2^6
25	x_0^1	x_3^2	x_2^3	x_3^4	x_1^5	x_0^6
26	x_0^1	x_5^2	x_4^3	x_0^4	x_0^5	x_4^6
27	x_6^1	x_1^2	x_0^3	x_3^4	x_1^5	x_0^6
28	x_2^1	x_0^2	x_0^3	x_1^4	x_0^5	x_2^6
29	x_6^1	x_6^2	x_0^3	x_3^4	x_0^5	x_4^6
30	x_0^1	x_4^2	x_4^3	x_0^4	x_3^5	x_6^6
31	x_4^1	x_5^2	x_0^3	x_0^4	x_6^5	x_3^6
32	x_0^1	x_2^2	x_3^3	x_0^4	x_0^5	x_1^6
33	x_6^1	x_0^2	x_2^3	x_1^4	x_4^5	x_0^6
34	x_6^1	x_0^2	x_0^3	x_2^4	x_2^5	x_3^6
35	x_4^1	x_3^2	x_0^3	x_0^4	x_0^5	x_5^6
36	x_5^1	x_0^2	x_6^3	x_0^4	x_3^5	x_6^6
37	x_0^1	x_4^2	x_5^3	x_4^4	x_3^5	x_0^6
38	x_0^1	x_5^2	x_6^3	x_2^4	x_0^5	x_1^6
39	x_0^1	x_6^2	x_2^3	x_1^4	x_5^5	x_0^6
40	x_1^1	x_0^2	x_0^3	x_3^4	x_0^5	x_2^6
41	x_1^1	x_0^2	x_5^3	x_0^4	x_1^5	x_6^6
42	x_0^1	x_6^2	x_0^3	x_4^4	x_0^5	x_6^6
43	x_0^1	x_3^2	x_3^3	x_1^4	x_2^5	x_0^6
44	x_3^1	x_0^2	x_3^3	x_5^4	x_1^5	x_0^6
45	x_4^1	x_2^2	x_0^3	x_2^4	x_6^5	x_0^6
46	x_1^1	x_0^2	x_5^3	x_0^4	x_4^5	x_3^6
47	x_5^1	x_0^2	x_0^3	x_6^4	x_2^5	x_4^6
48	x_0^1	x_1^2	x_0^3	x_5^4	x_0^5	x_1^6

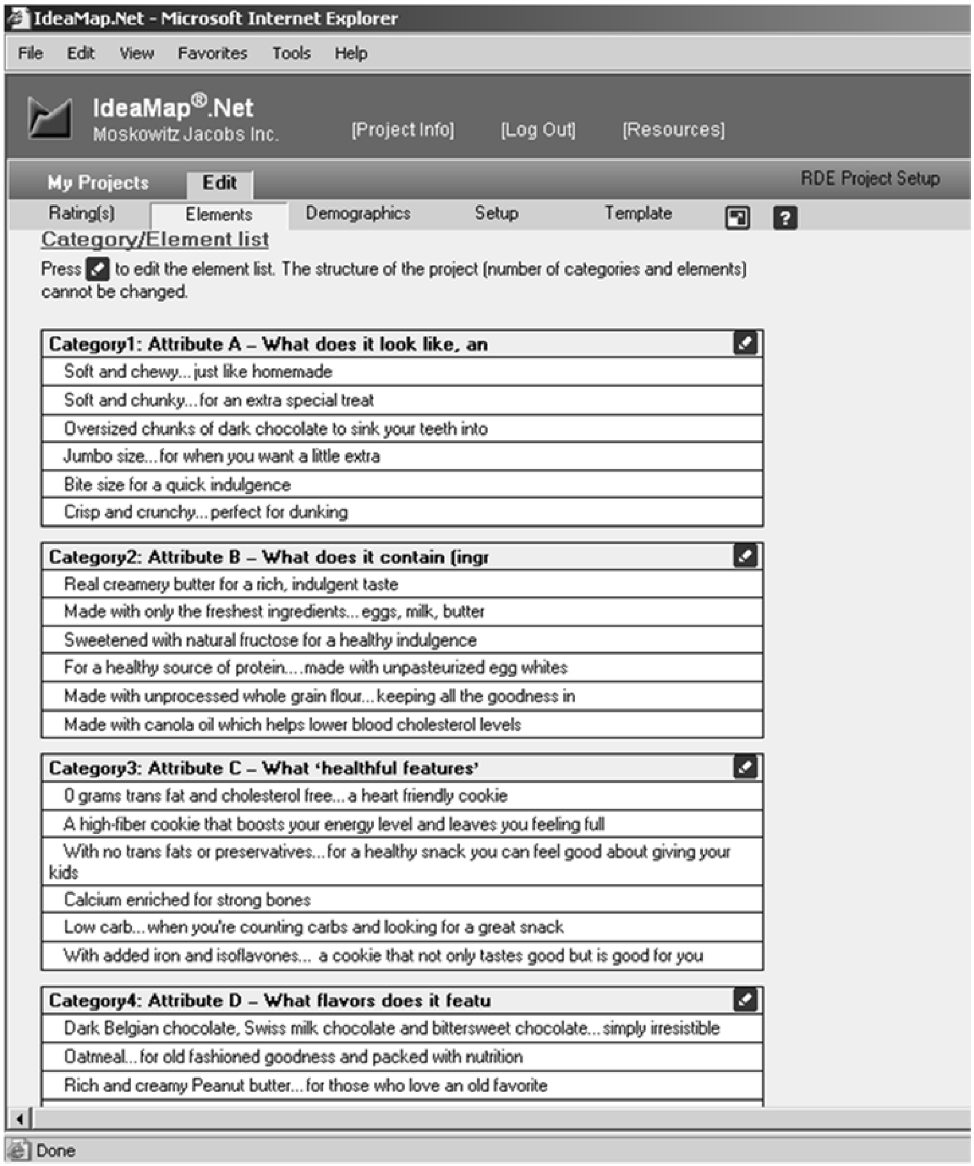


Figure 2: An example of the elements editing step in IdeaMap®.NET.

Step 3. Run the Experiment and Collect Data from the Consumers

In Step 3, the respondent evaluates individual sets of concepts. These concepts embody the experimental design. However, rather than seeing the structure of the design the respondent evaluates the combination of elements. The respondent assigns the appropriate subjective rating to each concept (Fig. **3**).

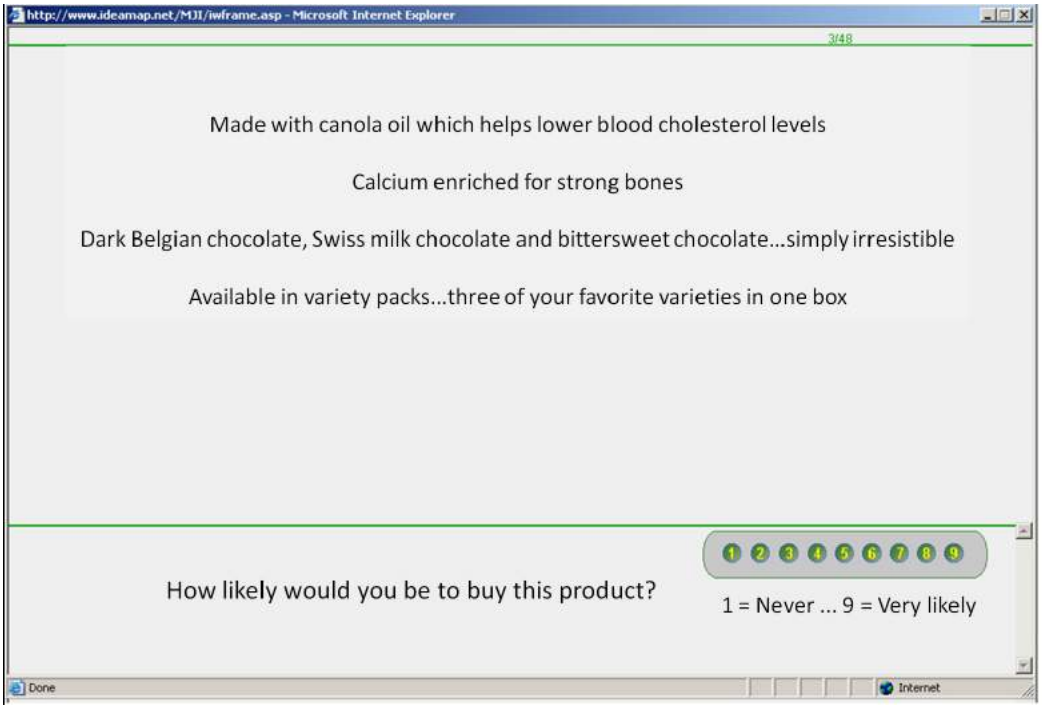


Figure 3: Screen capture of a concept evaluation page in IdeaMap®.NET. Every respondent evaluates a set of such concepts. The example shows all four attributes present in the concept.

Step 4. Analyze the Data to Generate Models or Relations among Variables

The ratings assigned by each respondent are analyzed using ordinary least squares regression. The elements are the independent variables; and the rating is the dependent variable. The regression model returns with an additive constant and an individual utility value for each element.

The key to using the model for product development is to understand the meaning of the utility score, *i.e.* the coefficient, for each element. To the degree that the element performs well by generating a high coefficient or utility score, the developer will be able to synthesize new and potentially better performing concepts. A combination of strong performing elements that were not hitherto combined may perform so well in the RDE study as to generate a breakthrough concept. On the other hand, to the degree that the concept elements perform only modestly, the concepts will not be a breakthrough. There are empirical norms for such performance based on thousands of previous experiments, most of which were conducted in the commercial realm (Moskowitz, Porretta and Silcher, 2005; see Table **3**). The numbers in Table **3** are conditional probabilities (%) of people who are interested in buying the product described. In turn, the additive constant in conjoint analysis with incomplete concepts is the baseline interest of respondents in the idea of the product (without any elements present).

Table 3: Norms for the Additive Constant and the Utilities Moskowitz *et al.* (2005).

Additive constant	**Interpretation**
>60	Respondents are very predisposed to the product
50–60	About half of the respondents are very positive
40–50	Respondents accept the idea
30–40	The elements need to do the work
<30	The product is a commodity and the elements must do the work
Utility score	**Interpretation**
>15	The element performs exceptionally well, breaks through clutter
10–15	The element performs well
5–10	The element breaks through the clutter
0–5	The element is barely effective, if at all
<0	The element actively detracts from acceptance

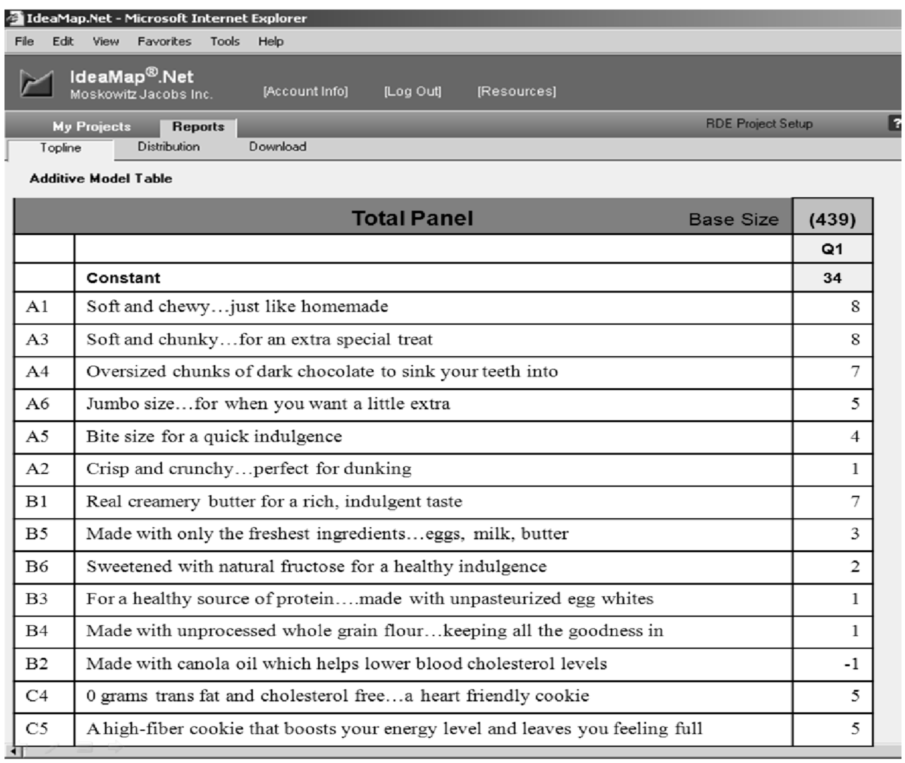
LOOKING DEEPER INTO THE MECHANICS

A great deal of understanding can be gained by looking at actual data in the way the computer program produces it. Knowing the format of the data, knowing what the results mean and spotting patterns right away constitute the hallmark of an individual who can move beyond the theory and into practice. With that in mind let us now look at the data as the computer program presents it. Our goal should be to mentally connect that which we do in the setup portion of the study with the computerized interview and then in turn with the results.

In our study, a total of 439 respondents completed the interview. The initial results of the study appear in Fig. **4**. Keep in mind that each of the 439 respondents evaluated 60 test concepts. Each respondent evaluated a set of concepts sufficient to create an individual-level model. Finally, the test concepts differed across respondents. Each individual generates an individual-level model. The utilities in Table **1** are actually averages of those individual-level models.

Let's move to the output of the IdeaMap®.Net as a researcher might see it. The RDE tool creates the individual-level model a few seconds after the respondent completes the interview and then stores it in a database. At any time the researcher can query the database to get "topline", *i.e.* the average utilities thus far. We see a screen capture of this topline in Fig. **4**. We actually captured that screen at the end of the study, so we have the data from all 439 respondents. Note that the elements are arrayed in descending order of utility. The reason for that is human nature; people who run the RDE experiments want to know what's winning, what's losing, but mainly what's winning.

We see the highest scoring elements in Fig. **4**. The additive constant is 34. We interpret that to mean that without any elements, approximately 34% of our respondents would rate the concepts as 7–9. Of course all concepts comprised elements, so our additive constant is really a computed parameter, a baseline.



	Total Panel Base Size	(439)
		Q1
	Constant	34
A1	Soft and chewy…just like homemade	8
A3	Soft and chunky…for an extra special treat	8
A4	Oversized chunks of dark chocolate to sink your teeth into	7
A6	Jumbo size…for when you want a little extra	5
A5	Bite size for a quick indulgence	4
A2	Crisp and crunchy…perfect for dunking	1
B1	Real creamery butter for a rich, indulgent taste	7
B5	Made with only the freshest ingredients…eggs, milk, butter	3
B6	Sweetened with natural fructose for a healthy indulgence	2
B3	For a healthy source of protein….made with unpasteurized egg whites	1
B4	Made with unprocessed whole grain flour…keeping all the goodness in	1
B2	Made with canola oil which helps lower blood cholesterol levels	-1
C4	0 grams trans fat and cholesterol free…a heart friendly cookie	5
C5	A high-fiber cookie that boosts your energy level and leaves you feeling full	5

Figure 4: Screen capture of Topline report page in IdeaMap®.NET. The numbers in the right column represent a conditional probability of consumers assigning top 3 rating scores on the rating question to the element.

Using the RDE Results to Create Winning New Ideas

We have been looking at RDE as a knowledge-producing system. When RDE is used by the scientific research community, its function is to expand our knowledge of a product or a situation. On the other hand, when a company uses RDE, the objective moves from simply learning to creating better concepts (Gofman, Bevolo and Moskowitz 2009).

To create a winning concept for the Total Panel (averaged across all respondents in the dataset), a developer should select the highest scoring elements from each of the attributes. Once the developer has selected the elements, it's simply a matter of adding up the separate utilities of the components that were selected, adding in the constant and then computing the sum. We see a simple example of such a summation below:

$$Ct_1 = A_1 + B_1 + C_4 + D_3 + E_3 + F_1 = 34 + 8 + 7 + 5 + 11 + 6 + 5 = 76.$$

We interpret this equation as follows: The estimated consumer interest (total utility score of the concept) is constructed as a sum of the corresponding additive constant and the utilities of each of the selected elements. Table **1** shows those utilities for all of the test elements. Continuing this train of logic, Ct_1 is the score of the optimized concept for the Total Panel; A_1 is element one of attribute A; B_1 is element one of attribute B, *etc.* The resulting sum (76%) is a conditional probability that the respondents would be interested in the product described by the selected elements.

From time to time, developers select combinations that they think will be acceptable to consumers. Such combinations of elements may not be as acceptable as they thought. Let's look at the case of a developer who wants to create the worst possible combination. If a developer by mistake selects the lowest scoring elements (*e.g.,* in the case when conjoint research has not been conducted and the selection was made based on predilections), the estimated consumer interest would be:

$$Ct_1' = A_2 + B_2 + C_3 + D_2 + E_2 + F_6 = 34 + 1 + -1 + 2 + -9 + 2 + 2 = 31,$$

where Ct_1' is the score of the concept comprising the lower-scored elements for the Total Panel. We learn from this simple exercise that the acceptance level can be doubled by experimentation. We also learn the lowest bounds of acceptance.

Knowing the limits of acceptance from the RDE data set has practical ramifications that would not be the case were the developer and researcher simply to test one concept at a time. By knowing the model that shows the utilities of the elements one learns immediately what is the best, what is the worst and thus both how high one can go and at the same time how low one can fall. Data need not generate a range from 76 to 31. It could turn out that all the utilities are small, meaning that the difference from high to low would be small, not large.

Step 5. Identify Pattern-Based Latent Segments and Interactions

Segmentation. The individual utility models can be used to create segments, as described above. These are pattern-based or response-based segments. The individuals in the different segments react to radically different ideas. The practical implication is that innovation for cookies would be best done first by specifying the particular segment to which one is going to appeal, then creating the optimal concept for that segment and afterwards creating the product to fit that concept. The example below demonstrates the power of pattern-based segmentation analysis.

The data suggest three clearly different segments with three different types of opportunities, only one of which (Health Seekers) offers any promise for major innovation (Table **4**):

1. *The Indulgents*: Segment 1, with approximately 66% of the respondents, looks only modestly promising for cookie innovation. This segment reacts very strongly to product features that we might call indulgent, such as *Oversized chunks of dark chocolate to sink your teeth into* (utility = 12) and *Dark Belgian chocolate, Swiss milk chocolate and bittersweet chocolate…simply irresistible* (utility = 13). This segment shows the highest among the segments additive constant (37), which is low enough to suggest that it will be the elements that do the work to achieve high concept performance. However, there are only a few elements that perform well, given the aforementioned norms for utility values. Simply looking at the proportion of respondents and recognizing this segment to be the most populous

will not answer the innovation issue because there is precious little to innovate for these individuals.

2. *The Health Seekers*: Segment 2 with approximately 13% of the respondents looks much more promising for cookie innovation. These respondents show a very low additive constant of 24, so the elements really have to work. Their interest is driven by many elements, including a few decadent flavor elements, but mostly health elements.

3. *The Traditional Treaters*: Segment 3 with approximately 20% of the respondents shows a modest additive constant (33), again suggesting that it will be the elements that do the work. The winning elements are fairly ordinary. These winning elements are *Soft and chewy…just like homemade* (utility = +14). *Soft and chunky…for an extra special treat* (utility = +14), *Oatmeal…for old fashioned goodness and packed with nutrition* (utility = +12) and *Rich and creamy peanut butter…for those who love an old favorite* (utility = +16). Looking at these promising elements we see a mirror of what is on the market. Segment 3 might be amenable to products with the aforementioned high-scoring features, but there is little innovation possible.

4. Despite the relatively low numbers of respondents it is clear that Segment 2, the Health Segment, shows the greatest promise for innovation. For this homogeneous-in-its-mind-set segment, the key elements are defined by the conjoint analysis; the elements that do well tend to be breakthrough; and it is possible to put together a strong selling proposition. Thus it is this group that should occupy the marketer's attention and it is this group that can be the target of breakthrough, innovative products, even in the mundane world of cookies.

Interactions. Knowing interactions can be very helpful in product development. When there are a few test elements, one can easily create the combinations using some type of "brute force" method, such as simply mixing the elements into pairs and testing the pairs as if they were single elements. Such brute force methods don't work well when we deal with 36 elements. In cases where we deal with six attributes and six silos per attributes, we have 540 pairwise interactions. That's simply too many to test in a simple RDE study, especially when we don't know which combination will interact, if any do at all.

Table 4: Element Utilities for the Total Panel and for the Three Concept Response Segments.

		Total	Seg1	Seg2	Seg3
Base size		**439**	**291**	**59**	**89**
Constant		**34**	**37**	**24**	**33**
Category A—What does it look like and feel like (appearance, texture)?					
A1	Soft and chewy…just like homemade	8	8	–2	14
A3	Soft and chunky…for an extra special treat	8	8	1	14
A4	Oversized chunks of dark chocolate to sink your teeth into	7	12	–3	–3
A6	Jumbo size…for when you want a little extra	5	5	0	6
A2	Crisp and crunchy…perfect for dunking	1	0	–5	6
Category B—What does it contain (ingredients)?					
B1	Real creamery butter for a rich, indulgent taste	7	8	17	–4
B5	Made with only the freshest ingredients…eggs, milk, butter	3	3	13	–4
B6	Sweetened with natural fructose for a healthy indulgence	2	–1	21	–3
B3	For a healthy source of protein….made with unpasteurized egg whites	1	–1	23	–9
B4	Made with unprocessed whole grain flour…keeping all the goodness in	1	3	12	–10
B2	Made with canola oil which helps lower blood cholesterol levels	–1	0	7	–9
Category C—What "healthful features" does it offer?					
C4	0 grams trans fat and cholesterol-free…a heart-friendly cookie	5	2	21	7
C5	A high-fiber cookie that boosts your energy level and leaves you feeling full	5	1	23	3

Note. Elements sorted by total panel (partial data). Seg. 1: Indulgent; Seg. 2: Health Seekers; Seg. 3: Traditional Treaters.

The experimental designs, especially the systematically permuted designs, allow the developer to identify combinations of product features, which exhibit unusually high or unusually low scores relative to what might be expected based on simple, noninteracting effects. Our cookie product example comprises six attributes of features and six elements per attribute. Doing the arithmetic brings us to 15 pairs of attributes, 36 interactions per pair, or a total of 540 interactions. That simple case cannot be dealt with using today's methods.

Let us now see how RDE handles these data. We described the method above, which first fits linear terms and then allows interaction terms to enter the model.

1. The combination of the elements A_6 (*Jumbo size…for when you want a little extra*) and E_6 (*Comes in a crush-proof box…no more broken cookies*) in a model that ignores interactions (a common approach) would produce neutral utilities of +2 for each of the elements, along with the additive constant +35. This could lead to a conclusion that it is not a very effective combination of features as the conditional probability of respondents being interested in buying this product is less than 40%. We get that value 40 by summing the additive constant (35) and the utilities for the two elements (2 + 2 = 4, 4 + 35 = 39).

2. A more precise model afforded by RDE with isomorphic permuted experimental designs (IPEDs) checks for all possible interaction between the elements of the project.

3. Using the foregoing approach by RDE (Step 2) above, we see that for Segment 2, only four pairs of elements show strong synergistic effects: $E_2 + F_2$, $A_6 + E_6$, $B_3 + F_3$ and $C_6 + E_3$ (Table **5**). In this case, the interactions effect is much bigger than utilities values themselves resulting in the total concept score (sum of the constant, individual utilities and interaction effect) of +59. This number suggests that there is a conditional probability of 59% that the respondents would consider buying this cookie product.

Table 5: How Accounting for Pairwise Synergistic Interactions Increases the Expected Concept Score, to Produce Better Performing Concepts.

	Pairs of elements that demonstrate synergistic effects			
	E2 + F2	**A6 + E6**	**B3 + F3**	**C6 + E3**
Element #1	Comes with a stay fresh inside wrapper… just twist and seal	Jumbo size… for when you want a little extra	For a healthy source of protein… made with un-pasteurized egg whites	With no trans fats or preservatives … for a healthy snack you can feel good about giving your kids
Element #2	In the frozen foods section of your supermarket…just thaw and serve	Comes in a crush-proof box…no more broken cookies	In the frozen foods section of the supermarket…bake and serve… hot from the oven	Comes in re-sealable bags. take out only what you want

Table 5: cont….

Results from the pure linear model (all respondents pooled into one regression)				
Additive Const.	35	35	35	35
Util. of El. #1	2	2	–6	2
Util. of El. #2	–3	2	–1	6
Sum	34	39	28	43
Results from the additive model which incorporates the pairwise interactions (all respondents pooled into one regression)				
Additive Const.	36	36	36	36
Util. of El. #1	0	1	–8	1
Util. of El. #2	–5	1	–2	5
Util. of pairwise interaction	18	21	19	14
Sum	49	59	45	56
Accounting for pairwise interactions generates this additional increase in concept acceptance				
Interaction effect	15	20	17	13

Note. Results taken from Segment 2 (Health Seekers).

Step 6. Apply the Generated Rules to Create New Products, Services, Offerings

Using RDE to "optimize" the ideas in the concepts leads to a set of quantitative relationships between the elements of RDE and consumer preferences and directions of how to increase consumer liking. The best combinations of elements are the ones that have the highest sum of utilities, including linear and interaction coefficients.

SUMMARY

The structured, disciplined, data-developing nature of RDE makes the process simple and streamlined. It leads to reliable results, both data and rules, that can be aggregated across studies, stored in the corporate vaults and IP (intellectual property) and then reused to answer new problems. The emergent quantitative relations between the elements include individual contribution of each element to consumer preferences as well as pattern-based latent segmentation and explicit and implicit interactions. Finally, as RDE generates absolute values of the utilities rather than relative values, it is possible to compare utilities across the attributes and even between some similarly structured projects conducted with the same type of consumers.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This chapter is partially based on Gofman, A. (2009). Experimentation-based product development in mature food categories: Advancing conjoint analysis approach. Tartu, Estonia: Tartu University Press.

REFERENCES

Gofman A. (2009). Experimentation-based product development in mature food categories: Advancing conjoint analysis approach. Tartu, Estonia: Tartu University Press (Doctorate Thesis).

Gofman, A. and Moskowitz, H. (2009). Steps towards a consumer-driven innovation machine for 'ordinary' product categories in their later lifecycle stages. International Journal of Technology Management, 45(3/4): 349–363.

Gofman, A., Bevolo, M. and Moskowitz, H. (2009). Role of corporate leadership and innovation claims in consumer perception of premium products. International Journal of Innovation Management, 13(4): 655–691.

Moskowitz, H. R., Porretta, S. and Silcher, M. (2005). Concept research in food product design and development. Ames, IA: Blackwell Professional Publishing.

CHAPTER 6

Consumer Metric Scales

Kannapon Lopetcharat[1,*] and Dulce Paredes[2]

[1]Nouveau Centric Co. Ltd., Bangkok, Thailand and [2]Takasago International Corp., USA

Abstract: The measurement of consumers' responses is the foundation of quantitative research. Selecting instrument(s) or scale(s) to measure responses is important and delicate, often confusing and politically sensitive. Researchers often find themselves in situations where they have to compromise but maintain the integrity of the study. To balance these tasks, the researchers' responsibility extends beyond simply designing a study. They must communicate and persuade other stakeholders in the project to accept their choices of scales. Knowledge of clearly established objectives and the inherent properties of the scales are critical. Research objectives dictate the choices of scales because the target consumers, questions and conditions for the study are derived from those objectives. In this chapter, the authors will recommend the appropriate use of scales in different situations (*e.g.,* screening/formulation-based experimental designs and discrete/final decision consumer studies). Besides clearly established objectives, the inherent properties of scales are very important in questionnaire design. The authors will present the latest development in hedonic measurements and the misuse of the 9-point hedonic scales in international studies. Moreover, the authors will review recent theories and evidence about ways in which "liking" and "disliking" are not diametrically opposed.

Keywords: Scales, hedonics, intensity, subjective measurement.

INTRODUCTION

The measurement of consumers' responses is the foundation of quantitative research. In practice, researchers often confront situations in which compromises must be made in choosing the metric to conform to standard or popular norms. To balance these tasks, the researchers' responsibility extends beyond designing the study to communicating and convincing stakeholders about what's appropriate.

Most rule developing experimentation (RDE) happens during early concept and/or product development. A lot of resources are usually devoted to designing and

***Address correspondence to Kannapon Lopetcharat:** Nouveau Centric Co. Ltd., Bangkok, Thailand; Tel: 845-667-4340; E-mail: info@nouveaucentric.com

Alex Gofman and Howard R. Moskowitz (Eds)

producing the appropriate stimuli or prototypes to represent the scope of the final product. It is equally important to understand the key consumer metric that best captures the most relevant consumer response that predicts optimal output for success. In some cases, several consumer metrics need to be optimized, creating tougher hurdles. Examples include identifying the most liked product that is also the most cost-effective to produce, identifying a winning formula in multiple markets, or identifying a winning formula that meets or surpasses internal norms or external benchmarks.

Several scales and their analyses have proliferated since Stevens (1946) first suggested four types of scales in measuring human perception: nominal, ordinal, interval and ratio scales. A substantial body of research compares relative performances of various scales to differentiate product stimuli and respondents' ability to use them. In this chapter we focus attention on common metrics used depending on the stage of product development, recent research on the popularly used 9-point hedonic scale and the latest development in hedonic measurements.

WHAT QUESTIONS SHOULD BE ASKED?

Research objectives dictate the test design and key metrics. The overall project objective in a new product process is to launch a winning product. The metrics change as the research evolves from initial to final product consumer testing in a systematic and iterative product research process. The metrics that matter should be chosen based on what key decision will impact the next stage.

During the most important phase of RDE, early development and screening, the attribute "overall liking" is often used as a key product. Overall liking is especially used when the respondents are to focus on the product itself, not on the brand, so the products are tested "blind", *i.e.* unbranded. The 9-point hedonic scale is the most popular acceptance measure, but researchers use other scales, such as the 3-, 5-, or 7-point hedonic scales. For work with children, simpler 3- and 5-point labeled or visual scales are commonly used. Overall liking metrics are usually carried through from initial to final product testing.

The 5-point purchase intent is the second most common metric especially for branded/concept testing since the percent top two box (percent definitely/probably will buy) is generally considered as a measure of repeat purchase. There is a need for other evaluative scales. Adding new metrics around ideal product performance, satisfaction and emotional connections during final product or positioning testing further differentiates a product winner in consumers' eyes. Practitioners are accustomed to developing scales and using them for various studies. These individually introduced scales are often idiosyncratic, being tailored to the needs and wants of individual researchers.

RATING SCALES FOR PARTICULAR STAGES OF DEVELOPMENT

Let's look at some of the ways that consumer responses are evaluated. Table **1** shows the different attributes that are used. Depending upon the stage, the rating scale will be different, because the respondent has to focus on relevant dimensions for that stage.

Table 1: Stages of Consumer Product Research and Commonly Used Metrics.

Stage	Purpose	Research Tools	Key Questions for Scaling
Concept development	Identify best features that are appealing to target consumers and/or key consumer segments	RDE message testing • Conjoint analysis • Concept screen	• Purchase interest • Overall liking • Uniqueness • Believability/relevancy
Early prototype development	Preference mapping or benchmarking to understand category aesthetic space for product optimization	RDE product testing • Category appraisal/sensory space design • Formula-based screening and optimization designs	• Overall liking • Purchase interest • Appropriateness for use • Sensory/aesthetics • Functionality or ease of use
Advanced product testing	Final screen to identify launch formula	• Acceptability test • Preference test (*vs.* control or benchmark)	• Overall liking • Preference • Purchase interest • Ideal attribute ratings (just about right)
	Claims substantiation	• Product use test (blind)	• Performance claims • Overall liking
	Concept product fit test	• Branded product use test	• Fit to concept • Purchase interest • Overall liking • Satisfaction • Emotional connection

The RDE works best when the objective is to discover patterns in data and create rules based upon those patterns. Patterns emerge with many stimuli, not with few stimuli. RDE thus fits best in the early development phase. The number of stimuli or variants decreases as the research progresses from early screening and optimization to final product validation. RDE studies early in the process identifies what drives the consumer to respond to the key question and efficiently screens out both poorly performing products and shows what attributes simply have no effect on acceptance.

The world of product evaluation provides RDE with number of well-established phenomena and, in turn, best practices. We list a few here:

1. RDE is based on responses to many products, not just one. Studies that use one product only and analyze responses to many attributes can't generate rules.

2. What is asked and how the questions are phrased play a key role. The sequence and position of questions should not be leading and should be focused on the product experience being measured.

3. The best data often come from respondents who test many products, not just one product. That is, seasoned RDE researchers would rather test many variants with fewer respondents during the early stage rather than test fewer variants with larger panels so as not to miss product opportunities.

4. A larger respondent base is used for final product testing and validation. At this stage, the final optimized formula is tested to validate that it meets key performance criteria against internal norms and in some cases, external benchmarks.

5. Performance is the key rating for many studies. Performance can be measured by different attributes. Both overall liking and purchase intent are used to assess overall product performance. A product that is highly liked will often receive high purchase intent ratings. However, there are certain products with functional characteristics where overall liking and purchase intent ratings diverge. A recent study, where both measures were asked for two types of beverages

(functional, indulgent) with systematically varied concepts, showed that correlation between overall liking and purchase intent was dependent on the particular stimulus. Overall liking and purchase intent correlated highly for the functional beverage, but did not correlate at all for the indulgent beverage. This research suggests that overall liking and purchase intent may operate differently and may need to become joint evaluative criteria and not mutual substitutes (Beckley, Moskowitz and Paredes, 2010).

6. Newer applications to areas beyond food work with attributes other than liking and purchase, because the "experience" is not just a hedonic or pleasure experience. For example, researchers working with functional products where hedonics (pleasure) is not the key criterion must use other metrics such as appropriateness for use, desired product benefit, or key distinguishing sensory perception to evaluate systematically designed or sensory differentiated products. For example, in a study of 18 eye make-up removers the researchers used comfort ratings from consumers and related it to descriptive data to create a perceptual map. They found consumers can differentiate a purely functional product based on perception of comfort (Delarue, Danzart and Siefferman, 2009).

7. The most popular key product indicator to understand choice and preference is still the hedonic measure of overall liking. However, overall liking does not measure the contextual aspect of the product such as its functional, emotional and social relevance. The impact of context and the corresponding consumer behavior is one of the hottest research topics today and is beyond the scope of this chapter. We focus here on hedonic scales only.

FUNDAMENTALS OF HEDONIC SCALES

9-Point Hedonic Scale (Historical Perspective)

One of the most popular instruments to measure consumer hedonic level toward stimuli is the 9-point hedonic scale. The scale was developed by the US military to measure differences in acceptance of various foods. Products that are well liked by consumers will score higher on the 9-point hedonic scale and products that are not as well liked will score lower (Jones, Peryam and Thrustone, 1955). The scale was first developed to make decisions about which foods to serve to soldiers (Lawless and Heymann, 1999). Fig. **1** shows the scale in one format, where the respondent simply checks the box that represents his opinion.

Figure 1: The 9-point hedonic scale (American English version) developed in the 1950s to gauge the level of food preference (Jones *et al.* 1955).

The history of the scale is worth a quick note because of the importance of creating a valid metric. The adjectives listed for the scale points (Fig. **1**) were selected from a list of 51 words and phrases. The adjectives were tested among ~900 soldiers. Jones *et al.* (1955) discussed the construction of the scale, which involved three steps:

1. The researchers selected a reference word, neither like nor dislike.

2. The soldiers assigned a numerical value from the following list (–4, –3, –2, –1, 0, 1, 2, 3, 4) to each of the 51 words.

3. The researchers rescaled the raw data for each word using Thurstonian modeling (Lawless and Heymann, 1999). The rescaling put the 51 words on a single scale. This strategy allowed the developers to select eight words, four phrases representing increasing levels of acceptance beyond the neutral middle and four phrases representing decreasing levels of acceptance below the neutral middle. Fig. **2** shows the actual Thurstonian scale values for the nine adjectives used in the 9-point hedonic scale.

The scale was constructed in order to assure two major properties desired in a measuring instrument: (1) interval property and (2) linearity. Each adjective in the 9-point hedonic scale was selected to have equal psychological distance

(Thurstonian scale value) from each other as possible. The linearity is a monotonic increase in hedonic magnitude from one adjective to the next. Both properties are "moderately" achieved through the development of the scale.

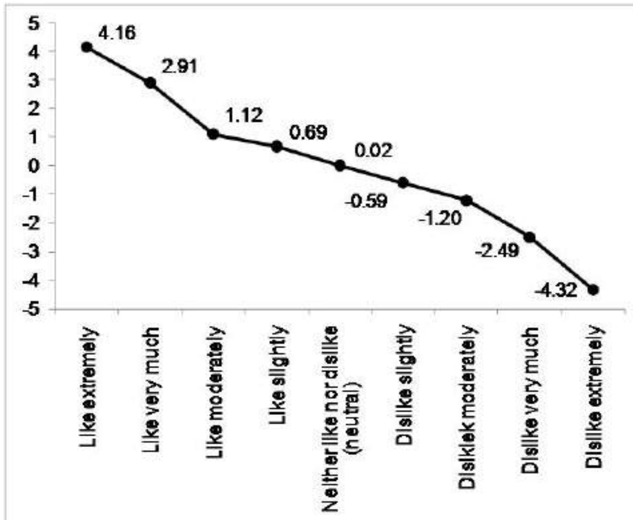


Figure 2: Thurstonian scale values for the nine adjectives used in the 9-point hedonic scale (Jones *et al.* 1955). Equidistance between the adjectives is not perfectly achieved; however, these were the most evenly distributed among the adjective tested.

PROBLEMS WITH THE 9-POINT HEDONIC SCALES IN CROSS-CULTURAL STUDIES

The original scale was "America-centric", created with soldiers from the US Army and employed by the US government for procuring foods and testing food acceptance. With the spread of consumer-centric research worldwide, it's becoming clearer that this scale may not work as well in other cultures:

1. **When and who:** Most of the subjects were American men in the 1950s and the adjectives selected were American English.

2. **Assumption and theory:** The assumption that liking and disliking are reciprocal.

3. **Applicability:** These terms may simply not translate well into other languages, for use with other cultures.

4. **Awareness:** Researchers are typically unaware of these problems. Most users of these scales are simply applying the scale, rather than thinking about the fundamentals underneath the scale.

The popularity of the 9-point scale among scientists and developers makes it worthy of additional attention in this chapter. The importance of the scale for RDE development cannot be sufficiently emphasized. And so we now look at the cultural differences, which abound in the use of scales. For example, Yeh *et al.* (1998) concluded that when literally translated versions of the 9-point hedonic scale are used, Asian consumers (Thais, Chinese and Koreans) use a smaller range (Yeh *et al.* 1998). However, the researchers did not confirm a major property of the translated scales: Consensual understanding of the words among target consumers. That is, are the meanings really the same? Is there consensus so that the respondents are really using the scale differently, or are they using different scales? The finding that *Asians use a small range of the scale* may have several causes. It could be the cultural tendency to avoid extremes. Or it could be caused by meaning confusion due to nonconsensual understanding of words used in the scales (the literally translated version of the 9-point hedonic scale in Thai, Chinese and Korean). Prescott (1998) reported that Japanese do use the extremity of the 9-point hedonic scale to express their opinion. One of the authors (Lopetcharat) has worked with Thai consumers and found the same to be true.

The consensual understanding of words is a fundamental property that directly governs the interval property and linearity of a scale. Even though the 9-point hedonic scale has been used successfully mostly in English-speaking cultures or among English-literate consumers, its two primary properties (interval property and linearity) have not been tested again since 1955, especially among other populations besides the American soldiers in 1950s. This problem is recognized by researchers but has never been widely rectified. The use of "literally translated" versions of the 9-point hedonic scale is prevalent in international studies just for the convenience of creating "norms" (Goldman, 2006; Prescott, 1998). Few studies reported this nonconsensual meaning phenomenon (Curia, Hough, Martinez and Margalef, 2001; Pedrero and Pangborn, 1989; Tuorila *et al.* 2009). Pedrero and Pangborn (1989) observed that Mexican consumers were confused when the 9-point hedonic scale was "literally" translated into Spanish. In

2001, Curia *et al.* reported that more than one-third of Argentines demonstrated reversion of meanings (ordering opposite from expected) when the scale was translated in Spanish. They recommended an unstructured scale instead of the 9-point hedonic scale as it was also recommended for Japanese (Prescott, 1998). Tuorila *et al.* (2009) reported no appropriate direct translation of the dislike portion of the 9-point hedonic scale in Finnish. One of the authors (Lopetcharat) experienced the same phenomenon when he tried to recreate the 9-point hedonic scale study in the Thai language and found many reversions in Thai literal translations, especially the ones in the middle of the scale. There appears to be only one publicly available report on non-English construction of hedonic scales (Daroub, Olabi and Toufeili, 2010). They repeated the development of an Arabic 9-point hedonic scale with extensive validation and mentioned no difficulty in translating English words to Arabic words.

Another issue in using the 9-point hedonic scale is the assumption that liking and disliking are opposite poles on the same dimension. In actuality, the 9-point hedonic scale forces a person to choose between liking and disliking before one can give a judgment toward a stimulus. Consequently, it allows a person to express only either liking or disliking. However, other studies demonstrate that a person can have both liking and disliking judgment simultaneously and the liking and disliking are not bipolar opposites on a continuum (Eagly and Chaiken, 1998; Herr and Page, 2004; Russell, 1979). Herr and Page (2004) reported empirical evidence to show that liking and disliking are *asymmetrically related as a bidirectional direction in memory* (Fig. **3**). In fact, the study demonstrated that liking and disliking influence each other through *priming* and disliking activates liking stronger than the liking judgment activates the disliking judgment. This phenomenon is called “negative dominant” and has been reported in many attitudinal and behavioral studies relating to good–bad or like–dislike (Rozin and Royzman, 2001).

In addition, Herr and Page (2004) suggested different mechanisms that govern the formation of liking (*automatic*) and disliking judgments (*controlled*). Automatic mechanism explains the spontaneous responses of liking and it contributes to *ability-related behavior* (*e.g.,* intelligence, skill, *etc.*). The controlled mechanism explains the slower responses of disliking and it contributes to *morally related behavior* (*e.g.,* honesty, concerns for other, *etc.*) (Skowronski and Carlston, 1989). However, the negative dominance did not exert its influence under extreme situations (*e.g.,* extremely like and extremely dislike something; Herr and Page, 2004). The impact of these two different mechanisms on the actual hedonic ratings was reported as different ideal-point structures were derived by using the bipolar 9-point hedonic scale, asking only liking and asking only disliking (the last two questions were asked under the halo effect of the 9-point scale (Drake, Lopetcharat and Drake, 2007).

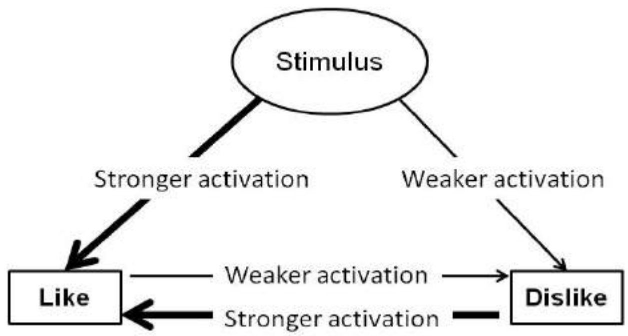


Figure 3: A structural relationship between liking and disliking judgment is simultaneously formed in a person toward a stimulus simultaneously before a final judgment is decided (to like or to dislike an object) (Herr and Page, 2004). The thickness of lines indicates the level of influences on each other (thicker line = stronger influence).

Therefore, in practice, the use of *bipolar* hedonic scales (*e.g.,* the 9-point hedonic scale or any of its variations) will inadvertently provide the illusion of the overall picture but, in fact, it gives a distorted and/or incomplete picture of consumers' attitudes toward a product (Olsen, 1999). Consequently, the measure of hedonics using these kinds of bipolar scales may not predict actual consumer behavior, as it has been proposed by many studies to use other measures such as appropriateness (Cardello and Schutz, 1996), purchase intent, choice-based judgment (Moskowitz, 1994), behavioral questions or indirect questions (Koster, 2003), emotional (Schifferstein, 2009), or two unipolar liking and disliking scales (Drake, Lopetcharat and Drake, 2009).

Your Preference for Product "A" over "B" Does Not Mean You Like Product "A" More Than "B": New Developments in the Hedonic Rating and Preference Tasks

When two products achieve equal liking scores, presumably one product should not really be preferred to another. Moreover, it makes sense to assume that when products become increasingly similar, the consumer's preference toward a product diminishes. Running preference tests after rating tasks, the so-called *protomonadic* test, is done based on an assumption that hedonic rating and preference judgment share the same underlying construct.

Research suggests that liking and preference are not the same (Simone and Pangborn, 1957). When asking consumers to make preference choices and when comparing corresponding hedonic ratings from the same consumers using the unipolar liking and unipolar disliking scale, Drake *et al.* (2007) found that, from 45 experiments, the level of disagreement between hedonic ratings and preference choices ranged from 30% to 80%. This apparent contradiction between liking and preference turns out to be independent of the magnitude of preference (Drake *et al.* 2007). That is, it was actually hard to predict preference from liking. The underlying mechanisms of preference and hedonic ratings are not the same, which may lead to different conclusions.

The practical implications for these findings could be critical for early stage development. The reason why a product is preferred over another product may not be the same as the reason why a product has a higher hedonic rating than another product. We are dealing here with different mental models of what is important, one mental model for liking, the other for preference. Such differences are critical, especially when the liking and preference tests generate insights about the product, with those insights becoming part of a company's intellectual capital.

NEW SCALING METHODS

Alternative hedonic scaling and measurements are constantly proposed to provide more suitable instruments for the practitioner to use in different circumstances (*e.g.,* cross-cultural study, different context of use, high carry-over effect of product class, *etc.*). In this section, we will provide helpful information to aid researchers in the selection of scales and brief explanation regarding the fundamental properties of alternative scales.

This section is organized into two subsections, *direct scaling* and *indirect scaling*, appropriate to the circumstances that surround products and consumers. For example, many products are used one at a time and rarely are compared directly to other products (*e.g.,* bubble bath, hair colorant, *etc.*). For these products, a monadic rating is more appropriate. However, many products can be compared in real life (*e.g.,* food products, fragrances, *etc.*). In this situation, pair-comparison tasks are more appropriate.

DIRECT SCALING

Direct scaling is a class of methods founded on the belief that consumers can express the magnitude of their liking (and/or disliking) by using a *scale* (a.k.a., *instrument* in social psychology and market research or *measurement* in information science, chemometrics, statistics, *etc.*; Lawless and Heymann, 1999). The expression of perceived hedonic magnitude can be captured using language (*e.g.,* the 9-point hedonic scale, categorical scales with anchors, *etc.*), using numbers directly such as the application of magnitude estimation for liking (Moskowitz and Sidel, 1971), or using semicross modality matching such as marking a tick on a line scale. The numerical values of the scales (except in magnitude estimation for liking) are assigned to the anchors or measured through physical measurement (*e.g.,* length of line, force exerted, distance of moving finger, *etc.*).

We focus here on alternative scales to measure acceptance. Recent development of alternative scales has attempted to overcome problems associated with the 9-point hedonic scale such as the elimination of problems due to translation and providing scale values that possess true interval-level information.

Labeled affective magnitude scale (LAM). Because the 9-point hedonic scale provides, at best, interval-level information (if one believes in the equal distance between the adjectives) or, at worst, ordinal-level information (if one conservatively ignores the claims of the equal distances), LAM was developed in 2000 (Schutz and Cardello, 2001) as an alternative to generate ratio-level hedonic measurement.

During the 1950s and parallel to the development of the 9-point hedonic scale, S.S. Stevens in Harvard University developed a scaling method that he ultimately called magnitude estimation (ME). According to Stevens, ME provides ratio scale values for perceived intensity of stimuli (Stevens, 1957). Despite its power and attraction to academic researchers, applying ME to practical problems in the world of hedonics is cumbersome. It is hard to perform correctly and hard to communicate to the research users in marketing and product development.

The ME generates numbers that represent intensities, yet there is no intrinsic and no practical meaning to these numbers. For example, what does it mean to have an ME score of 50 *versus* an ME score of 100? Does it mean that consumers like our products or not? ME needs a set of benchmarks, not so much for the statistics as for the interpretation, *etc.* Owing to the lack of these benchmarks, ME has not been commonly used.

The labeled affective magnitude scale (LAM) was developed to address this issue of benchmark. The LAM was developed to provide ratio-level hedonic information from consumers, marrying interpretive power with the simplicity of category.

Fig. **4** shows the scale and the logic. The position of the labels on the LAM was quantified by taking metric measurement along a 100-mm line scale and extensively validated against the 9-point hedonic scale (Schutz and Cardello, 2001). Schutz and Cardello proved that LAM provides ratio-level information and is as good as the 9-point hedonic scale in term of reliability, ease of use for individual consumers. LAM was found to be more discriminating than the 9-point hedonic scale when highly liked products were tested.

Figure 4: A labeled affective magnitude scale developed by Schutz and Cardello (2001).

Unipolar hedonic scale. A line scale with like and dislike anchored at two opposite ends without a middle mark or other scales presented in the same manner forces consumers to make a choice between like and dislike. The scale implicitly assumes that liking and disliking lie on opposite sides of one dimension. These forms of scales fall into a class of scale known as *semantic differential scales.* Those scales use two opposite words to anchor a scale (*e.g.,* line or categories). In practice and especially in the case of liking, many researchers simply assume such unidimensionality.

What happens when the scale is not really unidimensional but is treated as if it was? This can happen in the case of scales of liking/disliking. There are problems but also benefits to the unidimensional scale. Using two words that are not truly opposite confuses consumers. As a consequence, the results from the scale can be confounded. Many researchers notice this problem (Anonymous, 2009; Drake *et al.* 2007, 2009; Herr and Page, 2004; Simone and Pangborn, 1957) and proposed using a unipolar scale instead of a bipolar scale (or semantic differential scale) especially when truly opposite words have not been validated for the scale.

Potential advantages of a unipolar scale include:

1. Less work for consumers. They need to process only one word.
2. More likely to get useful information. Unipolar scales eliminate the untrue opposite meanings between two words.
3. In case of acceptance measurement, splitting like and dislike into two separate scales gives additional measurements for product success and failure.
4. There are instances where consumers use a liking scale (from No Opinion to Extremely Like) or a disliking scale (from No Opinion to Extremely Dislike) instead of rating a scale with liking and disliking on two opposite ends.

Rank-rating, positional relative rating (PRR) and rating with reevaluation and permission to change scores.

The rank-rating or PRR is an alternative to assess hedonic levels of products by combining two seemingly different tasks in one setting: (1) rating and (2) ranking.

The rating portion of this technique provides hedonic magnitude and the ranking reduces an error called *consistency error* (consumers rate two identical stimuli differently) or *reversal error* (an error made when consumer rates a stimulus with an actually lower intensity higher than another stimulus with an actually higher intensity or consumer rate) through the reevaluation (Koo and Kim, 2002; H.-J. Lee and Kim, 2001; O'Mahony, Park, Park and Kim, 2004). (Cordonnier and Delwiche, 2008) found that the reevaluation and permission to rescore (retesting a product in their paper) are the keys to increasing the discrimination power of rank rating.

INDIRECT SCALING

Indirect scaling refers to a class of methods that estimate the magnitude of hedonic responses using nonhedonic (behavioral, intent, appropriateness, *etc.*) or derived measures (regression coefficient, proportion, utilities, *etc.*). Indirect methods come from two beliefs about the ability of people to act as measuring instruments. The first one is based on Fechner's philosophy that consumers cannot express the magnitude of their perception directly (Fechner, 1966 (translation, orig. 1860)) and the second one is based on the idea that people are not really as reasonable as they think they are (Koster, 2003). Several indirect methods have been recommended as alternatives to direct hedonic ratings.

Asking indirect questions. People do not always make rational choices and are not always reasonable. In light of this nonrationality, Koster (2003) recommended indirect questions to gauge the level of liking toward a product or a stimulus. It is in the consumers' nature to try to figure out the purpose of the study and to please the researchers by answering the questions in a way that is thought to be socially acceptable (Koster, 2003). In short, consumers do what's expected. Normal people behave and don't analyze their experience; what they articulate is intellectualized at best. Therefore the strategy should be to ask the frequency of consumption (measuring hedonics) and then use some other method such as storytelling to measure the fit of product to a standard abstract idea such as authenticity, national food, or even formula alteration (Koster, 2003).

Simple pair-preference test. In a simple pair-preference test, consumers choose one of two products overall, rather than rating products on attributes, or rating differences. In short, the pair-preference test is a pair-comparison task coupled with a *preference* question. In a pair-comparison test, researchers usually have a rough idea that the *size* of the differences is very small. The goal is to confirm that the small differences are perceivable or not. In pair-preference test, the size of hedonic differences is the main question (Lawless, 2005). This fact differentiates the pair-preference test from the pair-comparison method and, consequently, it impacts the sample size, interpretation and conclusion of results. The result of a pair-preference test implies only the preference of one product over another product (Moskowitz, 2005).

In practice, many researchers use pair-preference tests to mimic an "actual situation" in the marketplace (Moskowitz, 2005). At the point-of-purchase, there are many brands for consumers to consider. In such cases the pair-preference test may mimic the actual situation. For example, consumers usually smell different shampoos or perfumes before they make a purchase decision. However, the inferences may not be relevant at the product-use level after purchase as consumers rarely compare products side by side or use two or more products at the same time. For example, when a regular consumer buys a new bubble bath product, the consumer will not have two bathtubs side by side or will not run between two different rooms to compare two bubble bath products (if the consumer actually has two identical bathtubs in the same house). Therefore, in this case, the rating task is closer to the actual situation (consumers use a product and make a judgment on how much he/she likes the product).

First–Last Choice Method (a.k.a. Max-Diff and Best–Worst Scaling). The First–Last Choice Method is a choice task. From a set of three or more options the respondent must select the best and the worst. This method has been used for more than 70 years by psychophysicists as a part of multidimensional scaling of perception (Richardson, 1938). It was later discussed by researchers in the 1950s as Torgerson's Method of Triad (three options with the choice of the two most different options; Torgerson, 1958). Recently, the First–Last Choice Method has been introduced to the sensory and product testing field under the name of Best–Worst Scaling (identifying the most liked and the least liked samples; Jaeger, Jørgensen, Aaslyng and Bredie, 2008) or Max-Diff Scaling (identifying the most and the least of a attribute) in the field of market research (Luce, 1959).

Therefore, Max-Diff Scaling and Best–Worse Scaling are special cases of the First–Last Choice Method (Ennis, 2009).

In the First–Last Choice Method, consumers are presented with a set of options and are asked to select two options from the set as an example below (Table **2**).

There are four (4) benefits or characteristics of cheddar cheese below. Which is the most and which is the *least* influential to you when you buy cheddar cheese?

Table 2: Example of the First–Last Choice Method.

The most	Benefit/characteristic	The least
X	Mild flavor	□
□	Made in New York	□
□	Organic	X
□	Low-fat	□

Note. Consumers select two mutually exclusive options from a set of four provided options.

After obtaining the frequency of consumers who select each sample/concept/statement as the least or the worst and the most or the best, the data can be analyzed in three different ways: Transforming to a probability scale by fitting a Multinomial Logit (MNL) model (Hein, Jaeger, Carr and Delahunty, 2008), converting the frequency to a Best–Worst Score (B-W score; Jaeger *et al.* 2008) or fitting a Thurstonian model (Ennis, 2009). B–W scores and MNL coefficients are related as B–W scores tended to be half of corresponding MNL coefficients (Hein *et al.* 2008). Hein *et al.* (2008) reported that the derived hedonic measure (B–W scores) is comparable to mean scores from direct hedonic scaling such as those from the 9-point hedonic scale and unstructured line scale. Best–Worst scaling is more discriminating and related to sensory attributes more than the responses obtained from the direct hedonic scales.

It is worth noting that the First–Last Choice Method, Max-Diff Scaling and Best–Worst Scaling are extensions of the simple pair-comparison or simple pair-preference method mentioned previously. Therefore, the results from this class of methods do not completely reflect the results from direct rating task (Drake *et al.* 2007). When the magnitude of differences (in any aspects of interest) is not big, the results may not agree between the two classes of methods. In addition, there are many disadvantages of using this class of methods, especially in product testing. The first is the test–retest task that consumers have to perform. The task restricts the method to only minimum carry-over effect products or products that are not time dependent. The second is the amount of time needed to conduct the study as it will take at least two-thirds longer than rating concepts and actual product testing (Cohen, 2003). Last but not least is the significantly more work needed to finish a study. For example, for six samples, this method requires the consumer to evaluate 30 samples.

Conjoint analysis. Conjoint analysis (CA) is a popular technique that researchers use to understand consumers' complex decision making, a process that consumers sense, perceive, assess, compare and/or evaluate stimuli (concepts, images, products, services, *etc.*) (Green and Srinivasan, 1978; Moskowitz, Beckley and Resurreccion, 2006, 2012; Orth and Lopetcharat, 2005). CA quantifies the magnitude of importance of stimulus attributes (also known as *elements*) using a derived measure from the analysis called *utility.* The use of the utility is the reason why the authors classify CA as an indirect method to measure hedonic levels of products.

Gathering the raw data for conjoint analysis uses both direct hedonic measures (*e.g.,* 9-point hedonic scale, LAM scale, line scale, *etc.*) and indirect hedonic measures (*e.g.,* pair-preference, choice, willingness to pay scale, purchase intent, *etc.*). The utility value is derived from the raw data through methods such as regression analysis. Conjoint analysis determines the so-called part-worth utility of each element of the product. Then the total utility of the product is estimated by summing the part-worth utilities of the different attributes. Conjoint analysis enables researchers to prioritize the attributes that were identified at the start of the project. Following such prioritization, the developer can select the appropriate strategy to create a new offering with these strong performing attributes.

Theoretical and technical mechanisms of CA have been extensively described (Louviere, 1988). In general, there are two classes of CA: indirect scaling CA and direct scaling CA (Moskowitz *et al.* 2006). Indirect scaling CA was developed in the 1960s and has been applied in market research since the 1970s (Green and Srinivasan, 1978) and it is still widely used today. Indirect scaling CA utilizes comparison task (pair-comparison, selecting from choice sets and ranking) as mentioned above (Louviere, 1988). A more recent development in CA is direct scaling CA, developed in the late 1980s and popularized by Howard Moskowitz (Moskowitz *et al.* 2006). Direct scaling CA is more efficient and provides deeper insights than indirect scaling CA because its utilities provide the magnitude of the hedonic response. The utilities derived from indirect scaling CA provide only the proportion of consumers (in percent) when the original data are counts from pair-comparison or choice tasks, or the probabilities of being in a rank order when the original data are rank orders.

One major advantage of CA, not so much statistical but in terms of process, is that CA forces researchers to think more thoroughly. CA demands the use of experimental design. Therefore, researchers must assess and reassess the research objectives before they can initiate the study. It is the process of systematic thinking, planning and implementing experimental designs and systematically exploring alternatives*,* that makes a CA study superior to other studies that do not involve experimental design (*e.g.,* a simple survey, monadic product testing, simple pair-preference test, *etc.*). Consequently, the results from a CA study, in general, are easier to interpret, understand and act on.

Willingness to pay and experimental auction. Willingness to pay (WTP) is another measure for gauging consumers' perceived values of products or services and it is believed to be a more accurate measure than hedonic and purchase intent measures when it is used in conjunction with experimental methods such as conjoint analysis or experimental auction (Lange, *et al.* 2002). It is expressed as a monetary value that consumers are willing to spend on a product, service, *etc.* Market transaction data are generally used to estimate WTP of existing products in the market. For new products, where a reference price does not exist in the market, at least three methods have been used to estimate WTP: (1) contingent valuation; (2) conjoint analysis; and (3) experimental auction (K.H. Lee and Hatcher, 2001; Rosen, 1974).

Contingent valuation (CV) is widely used to estimate WTP for products where comparable prices do not exist (Grunert *et al.* 2009). Its validity has been questioned since CV usually results in overestimation of the true WTP (Loomis *et* *al.* 1996). The aforementioned method of conjoint analysis (CA) is a widely popular method in market research to measure WTP. CA estimates WTP indirectly by adding *price as one of the attributes* in the study. With CA, WTP is expressed as the price utility instead of a direct expression of the amount of money that one is willing to pay.

The third indirect method is experimental auction (EA), which is closer to actual behavior because it makes the respondent bid on actual products. EA has been used widely in an experimental economy for estimating WTP. The most frequently used procedure is the Vickrey second price auction (Grunert *et al.* 2009), used to estimate estimating WTP for real goods and product concepts (Melton, Huffman and Shogren, 1996). The Vickrey second price auction works as follows:

1. Participants (bidders) are exposed to stimuli (products, concepts, brands, *etc*.). The amount of information that the bidders receive will help them to make the decision about WTP. Stimuli presentation is very important as it will dictate the inference of the WTP of the study.

2. After exposure to the stimuli, the participants submit their bids (their WTP) in sealed envelopes.

3. The bids for each stimulus are ranked from the lowest to the highest.

4. The winner (the participant who has submitted the highest bid) has the right to purchase the stimulus.

5. The price of the stimulus is set at the second highest bid (not the winner's bid).

6. Since the final price is not the highest price, the Vickrey second auction procedure induces participants to bid at their maximum WTP as the bidders know that he/she will pay less than his/her expectation.

CONCLUSIONS

Whether a researcher uses standard scales or the newer ones, the metric should be specific to the research objective and should produce accurate and actionable results. Regardless of stimuli tested (*i.e.* concept, message, actual product use), researchers must measure a response that is appropriate to the product in the relevant context. Understanding the fundamental properties of the response scales is critical to ensure a successful outcome of a consumer-based experiment.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Anonymous (2009). *Survey rating scale best practices*. Retrieved December 21, 2009, from http://www.vovici.com

Cardello, A. V., & Schutz, H. G. (1996). Food appropriateness measures as an adjunct to consumer preference/acceptability evaluation. *Food Quality and Preference, 7*(3/4), 239–249.

Cohen, S. T. (2003). *Research paper series: Maximum difference scaling: Improved measures of importance and preference for segmentation.* Sequim, WA: Sawtooth Software.

Curia, A. V., Hough, G., Martinez, M. C., & Margalef, M. I. (2001). How Argentine consumers understand the Spanish translation of the 9-point hedonic scale. *Food Quality and Preference, 12*, 217–221.

Cordonnier, S., & Delwiche, J. (2008). An alternative method for assessing liking: positional relative rating *versus* the 9-point hedonic scale. *Journal of Sensory Studies, 23*(2), 284.

Daroub, H., Olabi, A., & Toufeili, I. (2010). Designing and testing of an Arabic version of the hedonic scale for use in acceptability tests. *Food Quality and Preference, 21*, 33–43.

Delarue, J., Danzart, M., & Siefferman, J. (2009). More insights on consumer's perception of products – A case study on eye make up removers. Unpublished Poster presented at the 8th Pangborn Sensory Symposium. Firenze, Italy, July 26-30, 2007.

Drake, S., Lopetcharat, K., & Drake, M. (2007). *The role of dislike in food preference*. Paper presented at the IFT Annual Meeting. Chicago, July 28-August 1, 2007.

Drake, S., Lopetcharat, K., & Drake, M. (2009). *What does overall liking really mean: The effects of application and context.* Paper presented at the IFT Annual Meeting. Anaheim, CA, June 6-9, 2009.

Eagly, A. H., & Chaiken, S. (Eds.). (1998). *The handbook of social psychology*. Boston, MA: McGraw-Hill.

Ennis, D. M. (2009). Scaling Fist-Last, MaxDiff and Best-Worst data. [Technical report]. *IFPress, 12*(3), 2–3.

Fechner, G. T. (Ed.). (1966 (translation, orig. 1860)). *Elements of psychophysics*. New York, NY: Holt, Rienhart and Winston.

Goldman, A. (2006). Important considerations in cross-cultural sensory and consumer research – Introductory overview. *Food Quality and Preference, 17*, 646–649.

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5*, 103–123.

Grunert, K. G., Juhl, H. J., Esbjerg, L., Jensen, B. B., Bech-Larsen, T., Brunso, K., *et al.* (2009). Comparing methods for measuring consumer willingness to pay for a basic and an improved ready made soup product. *Food Quality and Preference, 20*, 607–619.

Hein, K. A., Jaeger, S. R., Carr, T. B., & Delahunty, C. M. (2008). Comparison of five common acceptance and preference methods. *Food Quality and Preference, 19*, 651–661.

Herr, P. M., & Page, C. M. (2004). Asymmetric association of liking and disliking judgments: So what's not to like? *Journal of Consumer Research, 30*, 588–601.

Jaeger, S. R., Jørgensen, A. S., Aaslyng, M. D., & Bredie, W. L. P. (2008). Best–worst scaling: An introduction and initial comparison with monadic rating for preference elicitation with food products. *Food Quality and Preference, 19*, 579–588.

Jones, L. V., Peryam, D. R., & Thurstone, L. L. (1955). Development of a scale for measuring soldier's food preferences. *Food Research, 20*, 512–520.

Koo, T. K., & Kim, K. O. (2002). Effects of forgetting on performance on various intensity scaling protocols: Magnitude estimation and labeled magnitude scale (Green scale). *Journal of Sensory Studies, 17*, 177–192.

Koster, E. P. (2003). The psychology of food choice: Some often encountered fallacies. *Food Quality and Preference, 14*, 359–373.

Lange, C., Martin, C., Chabanet, C., Combris, P., & Issanchou, S. (2002). Impact of the information provided to consumers on their willingness to pay for Champagne: Comparison with hedonic scores. *Food Quality and Preference, 13*, 597–608.

Lawless, H. T. (2005). Commentary: Every tool has its proper use. *Journal of Sensory Studies, 20*, 380–399.

Lawless, H. T., & Heymann, H. (1999). *Sensory evaluation of food: Principles and practices*. Gaithersburg, MD: Aspen Publishers, Inc.

Lee, H.-J., & Kim, K. O. (2001). Effects of forgetting on various protocols for category and line scales of intensity. *Journal of Sensory Studies*, 177–192.

Lee, K. H., & Hatcher, C. B. (2001). Willingness to pay for information: An analyst's guide. *Journal of Consumer Affairs, 35*, 120–140.

Loomis, J., Brown, T., Lucero, B., & Paterson, G. (1996). Improving validity experiments of contingent valuation methods: Results of efforts to reduce the disparity of hypothetical and actual willingness to pay. *Land Economics, 72*, 450–461.

Louviere, J. J. (1988). *Analyzing decision making: Metric conjoint analysis*. Newbury Park, CA: SAGE Publications.

Luce, R. D. (1959). *Individual choice behavior*. New York, NY: Wiley and Sons, Inc.

Melton, B. E., Huffman, W. E., & Shogren, J. F. (1996). Economic values of pork attributes: Hedonic price analysis of experimental auction data. *Review of Agricultural Economics, 18*, 613–627.

Moskowitz, H. R. (1994). *Food concepts and products. Just-in-time development*. Trumbull, CT: Food and Nutrition Press.

Moskowitz, H. R. (2005). Thoughts on subjective measurement, sensory metrics and usefulness of outcomes. *Journal of Sensory Studies, 20*, 347–362.

Moskowitz, H. R., Beckley, J. H., & Resurreccion, A. V. A. (2006). *Sensory and consumer research in food product design and development*. Ames, IA: Blackwell Publishing.

Moskowitz, H. R., Beckley, J. H., & Resurreccion, A. V. A. (2012). *Sensory and consumer research in food product design and development (2nd edition)*. Ames, IA: Blackwell Publishing.

Moskowitz, H. R., & Sidel, J. L. (1971). Magnitude and hedonic scales of food acceptability. *Journal of Food Science, 36*, 667–680.

O'Mahony, M., Park, H., Park, J. Y., & Kim, K. O. (2004). Comparison of the statistical analysis of hedonic data using analysis of variance and multiple comparison *versus* an R-index analysis of the ranked data. *Journal of Sensory Studies, 19*, 519–529.

Olsen, S. O. (1999). Strength and conflicting valence in the measurement of food attitudes and preferences. *Food Quality and Preference, 10*, 483–494.

Orth, U. R., & Lopetcharat, K. (2005). Consumer-based brand equity *versus* product-attribute utility: A comparative approach for craft beer. *Journal of Food Products Marketing, 11*(4), 77–90.

Beckley, J. H., Moskowitz, H. R. & Paredes, D. (2010). Beyond hedonics: Looking at purchase intent and emotions as criterion variables. Available at *http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1439733 (accessed on May 2, 2010).*

Pedrero, D. L., & Pangborn, R. M. (1989). *Evaluacion sensorial de los alimentos, metodos analiticos.* Mexico: Editorial Alhambra Mexicana.

Prescott, J. (1998). Comparisons of taste perceptions and preferences of Japanese and Australian consumers: Overview and implications for cross-cultural sensory research. *Food Quality and Preference, 9*(6), 393–402.

Richardson, M. W. (1938). Multidimensional psychophysics. *Psychological Bulletin, 35*, 659–660.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*, 34–55.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance and contagion. *Personality and Social Psychology Review, 5*(4), 296–320.

Russell, J. A. (1979). Affective space is bipolar. *Journal of Personality and Social Psychology, 37*, 345–356.

Schifferstein, H. N. J. (2009). The drinking experience: Cup or content? *Food Quality and Preference, 20*, 268–276.

Schutz, H. G., & Cardello, A. V. (2001). A Labeled Affective Magnitude (LAM) scale for assessing food liking/disliking. *Journal of Sensory Studies, 16*, 117–159.

Simone, M., & Pangborn, R. M. (1957). Consumer acceptance methodology: One *vs.* two samples. *Food Technology, 11*, 25–29.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105*, 131–142.

Stevens, S. S. (1946). On the theory of scales in measurement. *Science*, 103, 677–680.

Stevens, S. S. (1957). On the psychophysical law. *Psychology Review, 64*, 153–181.

Torgerson, W. S. (1958). *Theory and methods of scaling*. New York, NY: Wiley and Sons, Inc.

Tuorila, H., Huotilainen, A., Lahteenmaki, L., Ollila, S., Tuomi-Nurmi, S., & Urala, N. (2009). Comparison of affective rating scales and their relationship to variables reflecting food consumption. *Food Quality and Preference, 19*, 51–61.

Yeh, L. L., Kim, K. O., Chompreeda, P., Rimkeeree, H., Yau, N. J. N., & Lundahl, D. S. (1998). Comparison in use of the 9-point hedonic scale between Americans, Chinese, Koreans and Thai. *Food Quality and Preference, 9*(6), 413–419.

CHAPTER 7

Foundation of Sensory Optimization in the Food Industry

Howard R. Moskowitz*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The RDE applied to products involves experimental design, either of discrete or continuous variables. The chapter presents the history of RDE and sensory modeling/optimization, the methods most commonly used, field implementation, analysis procedures and segmentation. The chapter closes with a brief review of how the RDE approach is used by today's corporations.

Keywords: Sensory optimization, RSM, segmentation, product models.

INTRODUCTION

During the 1940s and 1950s, researchers in corporations began to realize that they need to better understand the link between formula/processing variables and responses. Companies such as Dupont in Wilmington, Delaware, were recognizing the value of systematic designs, where the chemists or product developers would combine several variables into a single mixture and then make the measurements on that mixture. The developer did not create one combination to be tested, but rather a number of different combinations, with the variables systematically changing. The combinations were created to be realistic alternatives to the end product, which itself was typically a mixture. The developer would test a number of different variables. Through simple statistics, such as averaging and regression analysis, it would become clear from these mixing experiments which variable(s) made a difference and how strongly each variable affected responses to the combination. And so was born the statistics, science and ultimately art of experimental design.

Known as response surface modeling (RSM), design of experiments (DOE) and often system modeling/optimization, the systematic approach would find a ready

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

audience in the world of product developers. There were a number of such advanced thinkers in the late 1950s and 1960s, especially Pillsbury Corporation's Dr. Al May (Joglekar and May, 1991) and Standard Brands' Dr. Robert Carbonell. The scientific literature of the time was also beginning to wake up to the power of such design, but there were relatively few published papers. Some of the early ones, published a mere 45 years ago (*e.g.,* Gordon, 1965) still resonate strongly because the power of experimental design transcends time.

WHO USES EXPERIMENTAL DESIGN IN PRODUCT DEVELOPMENT TO OPTIMIZE SENSORY REACTIONS?

Scientific methods do not spring fully formed and become immediately accepted by practitioners. Most advances take time to be proven, popularized and then adopted. It can take a quarter of a century or more from early appearance in the literature to the clear beginning of widespread adoption. Experimental design as the statistical method and rule developing experimentation (RDE) as the approach to learn about the product was no different.

Our story begins during the 1940s. During the decades between 1940 and 1970, most statistical treatment of sensory data involved inferential statistics: Did two products significantly differ from each other? Although such statistical analyses were popular and deemed the appropriate way to analyze consumer response to food, they did not teach the developer very much about how the consumer really perceived the food and how to change the food in order to increase acceptance. That latter information was deemed privileged knowledge, held closely in the mind of the product developer, not for venal reasons, but because there was no systematic body of knowledge based on experimental design.

The actual birth of RDE in the food industry can be traced to the early 1970s. Beginning in the 1970s, with the advent of widespread computer technology, especially the availability of time-sharing that put computation in the hands of many researchers, modeling of consumer and physical reactions to foods became increasingly popular.

The 1970s also saw another major change that would hasten the application of RDE in the food industry by bringing to the industries individuals who actually "thought" in RDE terms, although they did not realize it then. Psychophysicists entered the world of foods and beverages. Psychophysicists, experimental psychologists with a specialty in relating physical and perceptual worlds, thought in terms of models and equations, in terms of graphs and relations among variables. And so the growth of psychophysics and the study of how we process physical stimuli into taste, smell and texture sensations would find its natural outlet in the world of applied product development (see Stevens, 1975).

The first applications of experimental design to food processing were the aforementioned studies by Gordon and by May. It would take about 10–15 years before that pioneering work was recognized as a major sea change in product development. Beginning between 1975 and 1985, product developers in the US and Europe began to use experimental design in order to optimize the physical formulation of products, including foods (*e.g.,* salad dressings, meats, spaghetti sauces), beverages (colas, still beverages, coffee) and even fragrances (by changing the keys or the accords in the blend). What brought these changes about was the gradual acceptance by top management and marketing that a systematic approach to product development could shorten the development time and lead to many more successes (see Box, Hunter and Hunter, 1978; Cornell, 1981; Khuri and Cornell, 1987). It also helped that psychophysicists such as the author began to become heavily involved in commercially applying the principles of RDE to solve specific focused problems for many companies. By embedding RDE principles in the business practices of market research, the author introduced companies to the practical benefits of knowledge-based development through statistical design and modeling.

Today, the applications of sensory optimization techniques span the globe, enjoying success that can be traced directly to the early efforts of the 1970s. The small coterie of knowledgeable people in the 1960s and 1970s has expanded to included hundreds of professionals in the food industry. Part of the reason for the expansion has been the most obvious motivator of all: the approach works. Experimental design has helped products as diverse as Prego® pasta sauce, Vlasic® Pickles, Vanilla Dr. Pepper®, Equal® noncaloric sweetener, Hellman's® Salad Dressing and a host of other products achieve market leadership and increase the bottom line.

There is another reason, however—one more profound. Colleges are now teaching the principles of experimental design to students. Most schools that deal with product development, especially foods but also other products, teach the principles of experimental design. The result is what one might expect. The notion of systematically varying products, looking at an array of products, testing the different products and building mathematical models, is no longer foreign. It is not deemed a "waste of time" to build combinations of product features that are likely not to be acceptable. The naïve view that all the prototypes had to represent some level of "consumer acceptable product" has given way to the realization that it is better to build products that can fail in testing, so one knows the limits of formulations when it comes to produce products.

THE STEPS INVOLVED IN SENSORY OPTIMIZATION

As in any technology that grows in an organic way over time, sensory optimization does not comprise a simple, elegant structure, where you "follow the steps and get the answer". No, it doesn't work that way. There is a sequence of activities, not so much legislated, as proved to be productive over decades of use. We list the steps below in the following paragraphs, with some examples, some history and where relevant, with some observations, not so much about method, as about the nature of the specific method, where it comes from, why it is what it is.

1. Identify the specific variables that the experimenter can and will control in the study. This sounds fairly simple to anyone standing outside of the process. The assumption is often made that those working with the product or the process "know" what variables make a difference. The truth could hardly be more different. For the most part, companies don't do experimental designs to investigate the properties of their products. In turn, the companies remain unaware of the variables that drive responses, at least for most of their products. So, this first step of identifying the key variables is extremely important. One cannot "make up" in analysis for a missing, up-front, variable that is systematically manipulated by the experimenter.

2. Create the experimental layout. It is in this second step that statisticians should be involved. The experimental layout or test design specifies certain combinations of products that should be made. The combinations

are not random, but rather chosen to fulfill a number of requirements. Most important among these requirements is that the independent variables be statistically independent of each other. It takes a design crafted by the statistician or by any of a number of different computer programs (Box *et al.* 1978) to create such combinations. The second is that the design be able to accommodate some of the theoretical aspects. For example, if one is measuring degree of liking as dependent variable and one knows that liking may peak at some middle level, it is better to have three levels or more of the key variables, rather than two. Table **1** shows an example of an experimental design in three variables (A, B, C). This design, called the Box Behnken design (Box *et al.* 1978), has turned out to be among the most popular designs to identify optimal levels of products where liking peaks somewhere in the middle. The reason is simple. The three-level design is efficient, calling for very few prototypes, yet providing the ability to detect and reveal the nature of the nonlinearity between physical ingredient levels as the independent variable and liking as the dependent variable.

Table 1: Experimental Design (Box Behnken).

Prototype	**Variable *A***	**Variable *B***	**Variable *C***
1	1	1	1
2	1	1	–1
3	1	–1	1
4	1	–1	–1
5	–1	1	1
6	–1	1	–1
7	–1	–1	1
8	–1	–1	–1
9	1	0	0
10	–1	0	0
11	0	1	0
12	0	–1	0
13	0	0	1
14	0	0	–1
15	0	0	0

Note. The design shows the 15 combinations to be made for three variables, each that can appear at three "coded" levels (1 = High, 0 = Medium, –1 = Low).

3. Define the rating scales. Rating scales provide the language by which to relate the physical variables under the experimenter's control to the subjective responses under the consumer's control. With rating scales, the consumer becomes a measuring instrument. Of course, there is always a lot of variability among consumers in how they use the scale and even in the way they understand the terms on the scale. Table **2** presents examples of some rating scales.

Table 2: Example of Attributes and Rating Scales.

Liking: How much do you like this <product name> 0 = hate … 100 = love
Sensory: How SWEET is this <product name> 0 = not sweet at all … 100 = extremely sweet
Directional: How SWEET is this <product name> 0 = far too little sweetness … 50 = just right on sweetness … 100 = far too sweet
Image: Describe the nature of this product: 0 = only for children … 100 = only for adults

4. Run the "test". There are various ways to run RDE product tests. Many of the methods are covered by standard texts and articles about product research (*e.g.,* Griffin and Stauffer, 1990; Lawless and Heymann, 1998; Moskowitz, 1985; Stone and Sidel, 1985). Most of what is published covers the issue of good test practice, rather than RDE analysis, however. The key to a successful field execution lies in serving the products correctly, obtaining the appropriate rating data from each person for the product and ensuring that the products are rotated to reduce bias due to order. To the degree that a respondent evaluates many or most of the products over a single session, or multiple sessions across days, the data will become less "noisy". Each subject will act as his own control, often called a within-subjects design. The ballot requires the respondent to evaluate each product, one at a time, or so-called monadically (really sequentially monadically). This type of execution allows for easy data processing. The precise details and operational field considerations are beyond the scope of this chapter.

5. Prepare the data. Typically, data from the sensory optimization experiment generate a rectangular matrix, such as the matrix shown in

Table **3**. The products are the rows, the attributes are the columns. One may choose either to work with average data (one row per product), or with individual-level data (one row per product/respondent combination). The nature of the data set and the level of analysis (group *versus* individual) is left as a choice by the researcher.

Table 3: Example of Sensitivity Analysis (Independent Variable = Visible Herbs).

Vary: visible herbs	0	4	8	13	17	21	25
Cost of goods	788	795	801	805	807	807	806
Liking	61	63	64	66	66	66	66
Visible spices	16	30	41	50	58	63	66
Garlic aroma	53	57	60	61	62	63	62
Aroma strength	43	44	46	47	48	49	49
Flavor strength	55	57	59	60	61	61	61
Aftertaste	50	51	53	54	55	55	56
Saltiness	41	42	42	43	44	45	45

6. Create models using regression analysis. The models relate the independent variables to the ratings (see Draper and Smith, 1981). Each rating generates its own response. Modeling does not require a profound understanding of the product. Although many practitioners feel that one ought to know the dynamics of the product and how physical variables truly interact to drive a subjective sensory or hedonic (liking) response, that point of view is too stringent. The practitioner should, of course, make some assumptions about the product, such as the fact that as a physical stimulus increases, liking first increases, then peaks, then drops down. This assumption about the product means that in the regression model one should use both linear terms (X, Y) and quadratic terms (X^2, Y^2). At the end of the day, the majority of researchers who work with these product models choose simple linear or quadratic equations. The reason is simplicity; the equations are developed to be used in subsequent analysis and not as a more profound description of how the product actually works. An example of a polynomial equation appears below. The polynomial

shows, in schematic form, how two variables, A and B, in combination, "drive" the rating. The polynomial is estimated by least squares regression, available on many statistical programs:

$$\boldsymbol{Rating = k_0 + k_1(Variable\ A) + k_2(Variable\ A)^2 + k_3(Variable\ B) + k_4(Variable\ B)^2 + k_5(Variable\ A \times Variable\ B)}$$

7. Plot the equation. Although sensory optimization is, in its most profound essence, a mathematical approach, often the researchers like to graph the results, as we see in Figs. **1–3**. The figures themselves do not show as much as an equation can provide. Yet for many researchers a graph of some sort is emotionally more satisfying. There are three types of graphs: the three-dimensional plot of the actual (Fig. **1**); the sensitivity plot after a curve is fitted to the data and all points are brought to the curve (Fig. **2**); and the layer plot that shows contours, all points of which generate the same value for the dependent variable (Fig. **3**). Keep in mind when you look at these graphs that the graphs are "smoothed". That is, the regression analysis fits an equation. Afterward, the regression model estimates the likely rating for each combination of points. So, in the end, you see a smooth surface, rather than a bumpy surface. In fact, the great attractiveness of the graph is that is shows the general pattern, rather than forcing attention on the local bumps. We are more likely to spot patterns with these idealized representations of the empirical data.

8. Use the models to understand the dynamics of the product through "sensitivity" analysis. Sensitivity analysis estimates the value of the dependent variable given systematic changes in one variable. The other variables are held constant at some predetermined value. The results are clear both in tabular form and in graphical form. Fig. **4** shows an example of a graph. What becomes very important here is the actual learning, the shape of the curve. Such information is not typically known by companies, except those that do the product model. Sometimes the sensitivity analysis is not run by a curve, but rather actually estimated, as we see in Table **3**.

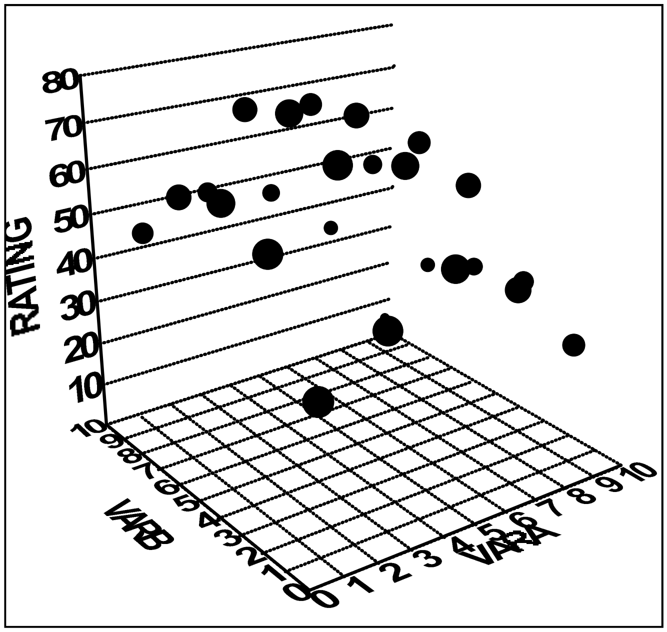


Figure 1: A plot of two variables (*A*, *B*) *versus* the rating. The plot locates each stimulus in the space. The size of the circles is proportional to the magnitude of the rating.

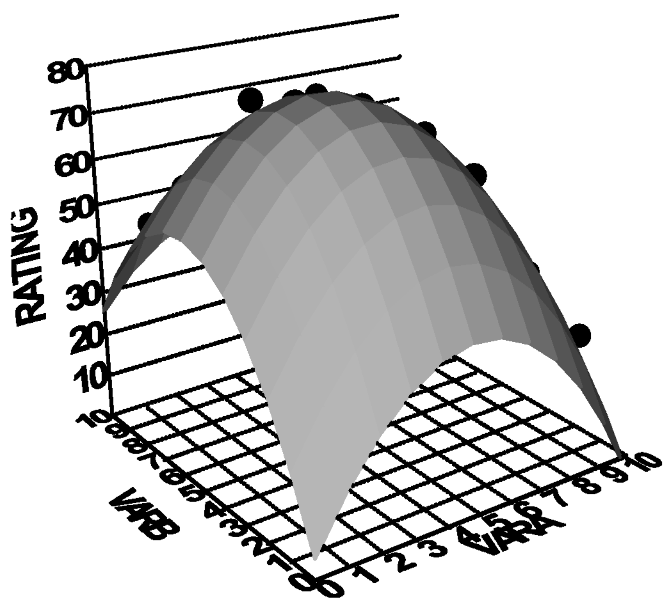


Figure 2: Smoothed surface. All points are brought to the surface by curve fitting. The smoothed surface shows the idealized pattern with regard to how two independent variables interact to drive the value of the dependent variable.

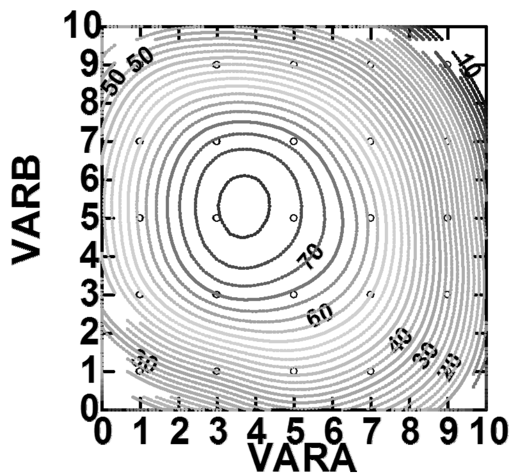


Figure 3: Contour plot. All pairs of independent variables (Var*A*, Var*B*) on a single contour with a given number combine to generate that level of the dependent variable.

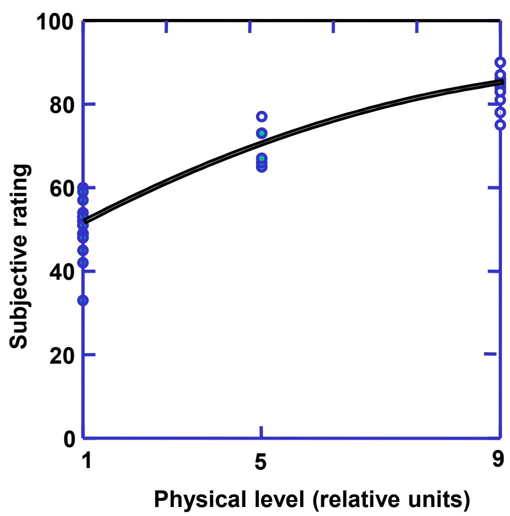


Figure 4: Sensitivity curve showing how liking changes with changes in the level of one formula variable. The curve is obtained from the product model, holding all variables but one fixed and estimating the liking (or other attribute) rating corresponding to the specific levels of the formula variable.

9. Optimize the product. The equation enables estimation of the likely rating for any combination of variables within the range tested, even estimations for combinations of variables that were not directly

evaluated. As long as the variables lie within the range tested, such interpolation is straightforward. That ability to interpolate is one of the greatest benefits of the product model. There are many different combinations, each of which has its own expected rating. One can discover the combination with the highest expected rating. This is called optimization. Table **4** shows an example of optimization. The optimization can be simply "what's the best product?" Or, when respondents have evaluated several attributes and if there are some objective functions such as cost, the optimization can search for "the best product within a specific cost" or "the best product that also satisfies constraints on the sensory profile". Applying RDE creates a corporate knowledge base and increases the ability to change the formulation to respond to current conditions, whether these are changes in the cost of goods (profitability), changes in sensory preference (marketing), or the desire to optimize the product so that it is less subject to quality defects (manufacture).

Table 4: Best Products, Showing the Effect of Imposing a Constraint on the Cost of Goods.

Maximize	**Total liking**	**Total liking**	**Total liking**
Constraint	None	Cost <500	Cost <470
Independent variables			
NaCl (salt)	100	100	100
Color	40	36	36
Visible herbs	19	19	19
Oregano	12	12	12
Basil	9	3	3
Pepper	40	32	32
Garlic	6	6	6
Thyme	3	3	3
Dependent variables			
Cost	570	500	470
Liking overall	74	73	73
Fit concept	70	68	68
Amount visible spices	63	64	64
Strength of aroma	38	38	39
Strength of taste	58	59	60

Table 4: cont…..

Juiciness	75	76	76
Italian flavor	61	61	61
Saltiness	40	39	40
Aftertaste	53	53	54

DISCUSSION AND CONCLUSIONS

What Do We Really Accomplish with Sensory Optimization?

This chapter deals with the design, testing, modeling and analysis of data to optimize the sensory characteristics of products. Yes, the methods are powerful, the statistics well-accepted, the analysis thorough. But what does this approach accomplish in the larger world of business? What are the key benefits of going through the systematic exploration and development of product "rules"?

We should explore two main areas. The first area is product design; the second area is competitive and financial advantage.

Sensory Optimization and Product Design

When marketers and product designers begin the process of creating a new product, typically they look at the consumer marketplace. Questions that emerge include what sensory characteristics seem to be associated with successful products, what types of "holes" exist in the marketplace that can be filled by new products and so forth. The questions are usually based on extensive research into the market as it currently stands, occasionally based on the "hunch" of smart marketers and very rarely on the "golden intuition" of one maverick individual, unless that individual runs the company.

The same type of disciplined analysis, instilled in business schools, cannot be said for product design and development, despite the availability of tools today. Developing the new product often proceeds by hit or miss, with developers creating prototypes, testing these prototypes informally (often at the bench where the products are developed, sometimes among colleagues or in a taste-test facility, sometimes in formal tests). The objective in the testing is to find out whether the product has promise and fulfills the marketing objective. Unfortunately, corporate culture and corporate funding force the developer to shorten the cycle, make do with best guesses, execute tests, but often use these tests simply to determine whether the product is reasonably acceptable. Today's development does not create the type of database that can be used as a corporate knowledge resource for years to come.

RDE-based sensory optimization provides a way out of the developer's dilemma. Instead of the randomized hit and miss, perhaps guided a little by the developer's intuition, sensory optimization proceeds in the structured manner described above. That is, the developer creates a variety of products, systematically varied, tests them and develops relations between the formula variables and the consumer response. Although the effort seems at first onerous, typically developers who go through the steps find that they learn a lot about the product. They can more quickly identify "what works" in terms of consumer acceptance, thus identifying the area of formulation that is most promising. The structure of systematic variations provides the matrix of alternatives that one can explore again and again, to identify specific product formulations that have the desired characteristics. The added benefit is the knowledge gained by systematic variation. This knowledge is invaluable and becomes part of the company's intellectual property.

Quite often the developer and by extension the corporation choose not to do this systematic variation, but rather opt to create the one or two "best shots". There are many compelling reasons not to do experiments. One consequence is that the development path may be quick when the first few prototypes just happen to lie near the optimal. The news is not so good when the prototypes don't deliver what is required. The development path becomes one of trial and error, fixing problems, only to find that new problems or shortcomings arise. There's no corpus of knowledge to guide one's effort.

A lot of trial and error, mainly error and lost opportunities may be avoided. Experimentation will help. It may be that no combinations can really deliver what is required. Experimental design with multiple products will show that immediately; no products in the design are sufficiently acceptable or have the appropriate sensory characteristics. That finding comes out immediately, because nothing "works" in the full set of 8, 10, 12 products, *etc.* It must be that this set of formula variables just won't deliver what's necessary. In contrast, a one-at-a-time approach can never reveal that. The developer, not knowing that it is impossible to deliver the product, continues to create and test prototype after prototype.

Managing the Product for Cost and Acceptance

Beyond the initial world of product development lays the world of managing the product during its lifetime. After the product is developed, but before it is launched, marketing and product developers typically analyze the "cost of goods". It does no one any good in business to deliver a superior product but lose money on every case delivered. And so the product that enters the marketplace always represents a compromise between the best that one can deliver to a happy consumer and the cost to do so. Few products ever survive when they are acceptable but their costs are too high to maintain. When such high-cost products enter the market and achieve success they must be reformulated anyway. The reformulation, designed to reduce cost, reduces the product quality right away. The product is then cost-optimized. And then the product will be withdrawn because it no longer sustains consumer interest. A product model will avoid all that.

Let's end this discussion of the product model with the cost issues, alluded to above. The traditional way to maintain product margins and thus profits in a time of changing costs is by testing different prototypes whose formulations are changed by the developer to reduce the cost of goods. The consumers who evaluate these cost-reduced (or occasionally quality-enhanced) prototypes do so with one or two prototypes. The feedback is either to accept the prototype as a prospective reformulation with lower cost, or to reject the prototype. When the latter unhappy event occurs, the developer goes back to the bench to create a new product.

With product modeling and sensory engineering, the sequence is quite different. Knowing the cost of goods of the systematically varied prototypes allows the developer to "dial a product". That is, the developer merely reestimates the cost of goods for each of the prototypes using "today's new ingredient costing", then reruns the optimization, looking for the combination of ingredients that will maximize acceptance, but entail a (new) cost of goods lower than a specified upper limit. The product model allows the company to manage the product for profitability, acceptance and delivery of the proper sensory experience.

Prospects for Sensory Optimization—Where Is It All Heading?

This chapter stresses the combination of science and business. Companies that use sensory optimization and experimental design do so knowing that the up-front work will entail effort and perhaps dislocations. Prototypes must be designed, made and tested. This is all work. The truth of the matter is that such efforts do take time; as a result, many companies avoid the work. Sometimes, as in the case of Prego®, Vlasic® Pickles, Maxwell House® Coffee, Vanilla Dr. Pepper® and many others, the opportunity to create a cost-effective, winning product is so great that the corporation accepts the challenge, assigns resources and in the end wins the prize: a better, more profitable product.

The foregoing is the positive. There is a downside. The downside is effort. Many companies today are risk averse, operate in a streamlined fashion and are committed to continuing with their past efforts. This is especially the case in the food and beverage industry. One consequence of current industry policies is that very few companies create "product models" using the principles of sensory optimization. There is too much effort involved and too much experimentation. Companies may conduct small tests of acceptance (so-called selection tests), or perhaps vary one or even two factors at a time in such small tests.

When a company is sufficiently enlightened to spend money on experiments and develop product models, the return may be tremendous. Such models would help the company manage margins, acceptance and sensory delivery for products for years to come. In the case of a number of companies with which the author has been associated for a decade or more, these product models are used for *years* to guide purchasing and reformulations. Using product models to guide decisions over a decade or more, with the model constructed after one RDE project, bears witness to the importance of sensory optimization as a knowledge-building, science and business tool.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Box, G. E. P., Hunter, J, and Hunter, S. (1978). *Statistics for experimenters*. New York, NY: John Wiley and Sons, Inc.

Cornell, J. A. (1981). *Experiments with mixtures*. New York, NY: John Wiley and Sons, Inc.

Draper, N. R., and Smith, H. (1981). *Applied regression analysis*. New York, NY: John Wiley and Sons.

Gordon, J. (1965). Evaluation of sugar-acid-sweetness relationships in orange juice by a response surface approach. *Journal of Food Science, 39,* 903–907.

Griffin, R., and Stauffer, L. (1990). Product optimization in central location testing and subsequent validation and calibration in home use testing. *Journal of Sensory Studies, 5,* 231–240.

Joglekar, A. M., and May, A. T. (1991). Product excellence through experimental design. In E. Graf and I. S. Saguy (Eds.), *Food product development* (pp. 211–230). New York, NY: Van Nostrand Reinhold – AVI.

Khuri, A. J., and Cornell, J. A. (1987). *Response surfaces*. New York, NY: Marcel Dekker Inc.

Lawless, H. T., and Heymann, H. (1998). *Sensory evaluation of food: Principles and practices.* New York, London, Tokio: Chapman and Hall.

Moskowitz, H. R. (1985) *New directions in product testing and sensory evaluation of food*. Westport, CT: Food and Nutrition Press.

Stevens, S. S. (1975). *Psychophysics: An introduction to its perceptual, neural and social prospects*. New York, NY: Wiley and Sons, Inc.

Stone, H., and Sidel, J. L. H. (1985). *Sensory evaluation practices*. New York, NY: Wiley and Sons, Inc.

PART II

PRACTICAL SENSORY OPTIMIZATION

CHAPTER 8

Introduction to Sensory Optimization

Howard R. Moskowitz*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The food and beverage industries are today facing an extremely competitive business situation. To the degree that the product developer or marketer, as well as general business manager, can understand the consumer and target efforts, the business will be more successful. This chapter introduces sensory optimization that could fill that need, for it provides both theory and case histories illustrating the types of issues, the nature of the thinking and the way the problem is solved in a practical format. Aimed toward all aspects of the industry, the chapter is especially important for those involved in the early stages of development, where there is much business opportunity.

Keywords: Sensory optimization, RDE, experimental design, sweetener, perfume, R&D, market research, industrial problem.

INTRODUCTION

Until fairly recently in the long sweep of history (say the first part of the 20th century), people ate what they could grow, kill and find. The notion of product acceptance as something to be measured was not part of the typical thinking of the citizen. Of course, there was always the good-tasting, the poor-tasting and the awful-tasting. One could not escape the fact that the natural course of events for the chemical senses, taste and smell, was to produce good- and poor-"tasting" stimuli. But the reality was that unless a food or beverage tasted bad or was spoiled, it would get consumed.

The agricultural revolution, the development of an aggressive food industry oriented toward profits and the increasing mechanization of food growing, harvesting and processing changed all that. From being happy to go to sleep on a full stomach, consumers developed discriminating palates. The change didn't happen overnight, of course, but it did happen.

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

And so our story begins: not so much at the basic level of taste and smell, or chemosensory processes, but where industry meets the palate and the pocketbook, at the level of the food and beverage itself.

PRODUCT TESTING

The precursor to rule developing experimentation (RDE), sensory optimization came in the form of tests about the product. You have to be a bit of a historian to dig back almost a century to the 1920s and a bit later to the 1930s and 1940s (Moskowitz, 1983; Stone and Sidel, 2004). It was becoming increasingly clear that foods were cheaper, that quality was beginning to achieve some level of attention and that the food industry was responsive, if not yet particularly knowledgeable. The notion of systematic experimentation had not yet penetrated the food processing world, although it seemed to be taking the world of agronomy by storm. Agricultural economics and agronomists were doing "planned experiments": planting fields with different seeds, fertilizing to different amounts and then measuring the yield.

At the same time, our processed food industry was progressing down a bit of a different path. In a moment, the difference in paths will become obvious. Agronomists and agricultural engineers are interested in the yield. Very simply, how much goes into the ground and of course what comes out. It's pretty simple. Farmers and agronomists know how much seed they use and what they harvest. They've lived with that for hundreds, nay thousands, of years. The food industry was following an entirely different path. The food industry wasn't focusing on yield from processing. When it came to food, the food industry needed to measure something more wraith-like, more elusive. This was the subjective response to the food. And so RDE and other systematics would have to wait.

Replacing systematized product testing, or better, antedating it and preparing the way, was product evaluation. The early research divides into two distinct parts:

1. Creating instruments that could simulate the chemical and textural aspects of food. This was a good half or more of the effort. Read the *Journal of Food Science* from the 1930s to get a sense of the types of machines that researchers used to assess the quality of foods.

2. Creating questionnaires that probed people for the reason that they liked a food. The researchers of those years were not sophisticated as we are today. It's not that they were less smart. Rather, it's that they did not have decades of thinking and research about the way to understand what consumers want.

At the end of the day, however, we can look back at those years and see the beginning of attempts to understand "what makes foods work" in terms of consumer acceptance. It is the 1960s to which we must turn now for our history of sensory optimization and its real contributions.

THE 1960S: ERA OF DESIGN, STATISTICS AND PROFESSIONALIZATION

If the 1930s and 1940s are to be considered the ancestors of RDE, then we should assign the 1960s to be the early childhood of the field. By 1960, the notion that the consumer could act as a measuring instrument was fully accepted by the industry (Jacoby, 1978). The Institute of Food Technologists (IFT) in the United States, the organization encompassing many of the food researchers, began holding yearly, ever-bigger conventions in different cities, bringing together researchers and practitioners. In these IFT meetings, various sessions were devoted to the systematic approach to increasing food acceptance. RDE and sensory optimization were not yet being thought of as a separate discipline, but when you read the abstracts from those meetings you get a sense that practitioners knew that systematic approaches to optimization were necessary.

Sensory optimization appears to have gotten its start in a number of different places in the 1960s, although as we will see, various professionals claim that they were doing it back in the 1950s. It really doesn't matter when the first studies were being done. What is important is when the field took off.

In the 1960s, statisticians were advocating systematic design of experiments. The notion of such design is, as we have seen, not particularly new; agronomists were using designed experiments in the 1920s through the 1940s to increase crop yield (Yates, 1964). But during the 1940s, statisticians working on experimental designs for military purposes (*e.g.,* Plackett and Burman, 1946; Rao, 1947) were released to work for industry after the war's successful conclusion in 1945. And work they did, with a proliferation of unpublished corporate papers on applications.

Our story jumps to the 1960s. The American Society for Testing and Materials (ASTM), headquartered at that time in Philadelphia and noted for standards in the construction and chemical industries, founded Committee E-18, called Sensory Evaluation (Hootman, 1992). This committee was charged with developing standards for evaluating products. But the committee did more. It was the first committee that formally recognized the developing role of sensory analysis in the evaluation of products, especially foodstuffs. And it would be Committee E-18 that would become the center point for efforts on sensory optimization. All those involved in optimization, whether statisticians or researchers applying the statistical principles, had somewhere to go. Representatives from DuPont Chemical, such as Mary Whitcomb Jenkins, experts on experimental design as taught in DuPont, would present the approaches during the committee meetings. And the field took off from there.

ENTER PSYCHOPHYSICS

The growth of a scientific discipline is never as simple as one discovery leading to another, one world view leading to the other and then displacing it. And the same is true for sensory optimization. It is not really the case that agronomy led to designed experiments and that somehow product developers picked up on this track and adopted designed experiments for products. The story is a bit more convoluted than that.

In the 1940s, experimental psychologists in a number of laboratories spread around the world were working on the measurement of subjective perception. Specifically, interest at the time focused on the person as a measuring instrument. Of course, to many of us in end of the first decade of the twenty-first century, such a research focus seems a bit pointless. Our daily life is suffused with scales. Everyone is accustomed to measuring something or other, usually on a computer and typically having to do with one's experience with a website, a transaction and so forth.

This was not the case 70 years ago. At that time, experimental psychologists were struggling with the best way to have people measure their perceptions. There were all sorts of arcane ways, such as response time or degree of preference (Rockett, 1956). In the end, however, it would be simple rating scales that worked.

Those efforts of experimental psychologists would influence the food industry and ultimately lead to RDE and to sensory optimization. Once psychologists got over the notion that subjective measurement of liking and sensory experience was difficult, arcane and impossible almost because people were filled with errors, the field "took off". Researchers began to measure food acceptance and sensory perceptions, using simple scales (*e.g.,* 1–5, 1–9, 1–100) and even more powerful but harder to implement scales such as magnitude estimation (Stevens, 1975). The bottom line, however, was that experimental psychologists "accepted and blessed" the use of people as measuring instruments.

With these efforts came the attempt to relate subjective responses to physical stimuli. Notable among these efforts were the pioneering papers of S.S. Stevens at Harvard University. After literally dozens of experiments, Stevens reported that there seemed to be a recurring relation relating subjectively rated intensity (S) to physical stimulus intensity (I). The equation best fitting that relation was a power function of the form: $S = kI^n$. The exponent, n, was repeatable from study to study and varied from a low around 0.2 for odor, to around 1.0 for subjective intensity of length, to a high of 2 or more for the pain felt for shock (see Table **1**). Fig. **1** shows an example of these relations, which Stevens reported when he instructed panelists to rate sensory intensity.

Table 1: Exponents for the Power Function $S=kI^n$, Relating Subjective Intensity S to the Physical Magnitude I.

Continuum	Measured Exponent	Stimulus Condition
Loudness	0.67	Sound pressure of 3000-Hz tone
Vibration	0.95	Amplitude of 60 Hz on finger
Vibration	0.60	Amplitude of 250 Hz on finger
Brightness	0.33	5° target in dark
Brightness	0.50	Point source
Brightness	1	Point source briefly flashed
Lightness	1.2	Reflectance of gray papers

Table 1: cont….

Visual length	1	Projected line
Visual area	0.7	Projected square
Redness (saturation)	1.7	Red-gray mixture
Taste	1.3	Sucrose
Taste	1.4	Salt
Taste	0.8	Saccharin
Smell	0.6	Heptane
Cold	1	Metal contact on arm
Warmth	1.6	Metal contact on arm
Warmth	1.3	Irradiation of skin, small area
Warmth	0.7	Irradiation of skin, large area
Discomfort, cold	1.7	Whole body irradiation
Discomfort, warm	0.7	Whole body irradiation
Thermal pain	1	Radiant heat on skin
Tactual roughness	1.5	Rubbing emery cloths
Tactual hardness	0.8	Squeezing rubber
Finger span	1.3	Thickness of blocks
Pressure on palm	1.1	Static force on skin
Muscle force	1.7	Static contractions
Heaviness	1.45	Lifted weights
Viscosity	0.42	Stirring silicone fluids
Electric shock	3.5	Current through fingers
Vocal effort	1.1	Vocal sound pressure
Angular acceleration	1.4	5-sec rotation
Duration	1.1	White noise stimuli

Note. The exponent n is the key parameter of interest.

SENSORY OPTIMIZATION TAKES OFF

The world of product research assimilated the joint contributions of statistics on the one hand and psychophysics on the other. The statistical approach provided efficient experimental designs, allowing the developer to work with fewer than the very large number of combinations that might be required when working with four, five, six, or more variables. For example, with six variables, each at three levels, the full set of combinations would entail 3^6 combinations or 729 combinations. No sane researcher would or could spend the necessary time creating the requisite number of combinations and then testing them. The effort would simply be too great. In contrast, statisticians have created experimental designs that make such studies very easy (Box, Hunter and Hunter, 1978). For example, by using the Plackett–Burman, 3-level screening design, one needs a mere 27 combinations, well within the level of effort that a company would invest for an important product.

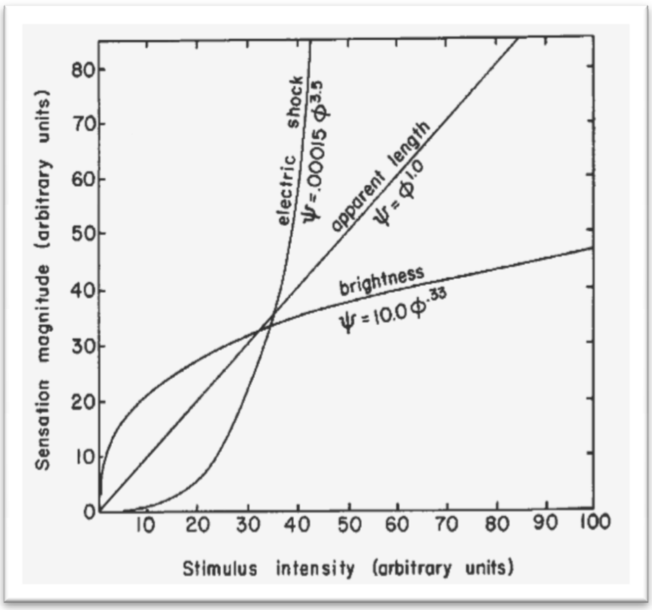


Figure 1: The schematic relation between physical intensity (abscissa) and subjectively perceived intensity (subjective magnitude, ordinate). The curves can be described by power functions, of the general form: $S = kI^n$.

Psychophysics, in turn, provided a different gift to sensory optimization. Psychophysicists entered the food industry and their basic interest in sensory perception to understand how we perceive the different characteristics of food. Psychophysicists adapted the methods they used to study sensory process from the laboratories of basic science and used these powerful methods on more applied problems. These researchers brought from academia solid approaches to understand sensory processes, but now instead of using science to figure out how the senses worked, they were using science to understand the product. It was the same discipline; only the focus differed. Examples of approaches used by psychophysics and applied to the world of industry-based development appear in five books by the author, beginning in the early 1980s and stretching over a 13-year period (Moskowitz, 1983, 1984, 1985, 1994, 1996).

SOME EARLY STORIES FROM THOSE DAYS

Stories make research and science come alive. Sensory optimization becomes even more interesting when we can see it in action. We will talk about three of the studies, but not so much in detail as examples of what can be achieved. These will be the story of Equal® (aspartame), the story of Kotex® napkins and finally the story of Chesebrough Ponds and the search for an optimal fragrance for Brazil. There are different lessons to be learned.

RDE AND THE EARLY DAYS OF SWEETENER EQUAL®

During the early and mid-1970s, a frantic search was on for new sweeteners. The low-calorie sweeteners of those days were saccharin, which had been around about 100 years and cyclamate, which had been around since the 1930s. Government and private research continued to report problem with sweeteners when fed to rats in high doses. At the time rumors were swirling around that one of the two sweeteners, cyclamate, was in trouble.

In Chicago, meanwhile, Fermco Biochemics, a small company specializing in yeast products, had discovered a sweetener called aspartylphenylalanine methyl ester (aspartame) (Cloninger and Baldwin, 1970). Headed by Dr. Don Scott, Fermco began the long, arduous process of getting aspartame accepted as a high-potency sweetener.

As part of the effort, Fermco sponsored a large-scale RDE study, mixing together different combinations of the sweeteners aspartame, calcium cyclamate and sodium saccharin. The different combinations were developed according to an experimental design and put together by RC Cola® under the guidance of its technical director, Dr. Martha Jones.

The studies were run, the data collected from a group of respondents who evaluated the different products for liking and sweetness and the results analyzed. What is important in the story is that Fermco used the data for the FDA-required process in order to get the sweetener accepted. And, to make it more relevant, the resultant equations that RDE generated were used to identify specific combinations and estimate the mixture sweetness of the combinations.

As part of the vetting process for aspartame, the results of this RDE study were published in a reputable scientific journal, over 30 years ago (Moskowitz, Wolfe and Beck, 1979). Standing back, we see from the distance of a third of a century the promise to the food industry that RDE and sensory optimization were poised to deliver.

OPTIMIZING THE DIMENSIONS OF SANITARY NAPKINS

The original use of RDE in industry was to deal with food and beverage formulations. In actuality, however, RDE scored one of its most interesting triumphs as a tool by which to learn about comfort and protection.

Our story takes us to Neenah, Wisconsin, in the late 1970s, during the time when RDE was fighting to be accepted as a research procedure in areas outside the world of food and drink. At that time, Dr. Elaine Jeveli of the Kimberly Clark Corporation recognized that with RDE she could begin to systematize the company's understanding of what comfort was about. Not that systematic exploration of sanitary napkins would be the "be-all and end-all" of comfort, but rather with systematic variations of the physical dimensions (length, width and depth) she could understand what factors drove users' feelings of comfort, protection and discreteness.

Setting up the RDE study for sanitary napkins was both easier and harder than had been imagined by the researchers. For instance, the scope of the study had to be considered: Women would have to wear the product, rather than experiencing it for just a moment in a typical "taste test". Thus, the physical levels were to not only be realistic, but also push the current limits. That was the first problem to be solved; not particularly difficult but requiring some thought.

The more difficult problem was how to run the study. The actual RDE study called for three variables (length, width and thickness), each at three levels (somewhat greater than current, the same as current and somewhat less than current). With this design, the worst case was 27 prototypes, out of consideration for the practical world of business. Experimental designs, however, reduced these 27 combinations to a more manageable 15. Still, a woman would wear a sanitary napkin for five days maximum. So, the typical psychophysical approach would not work, where everyone evaluated every product. Even more vexing was the fact that the woman was "not the same" across the five days, with some days defined as "light flow" and others defined as "heavy flow".

The actual issues shouldn't concern us here; just the story. The bottom line was that the study was run effectively: Each woman evaluated three products on the three heaviness days, one product per day, rated the products and in two months the study was finished. The bottom line: a great deal of learning about how physical variables drive actual protection, perceived protection, perceived comfort and perceived discreteness, a first in the industry.

RDE AND THE CREATION OF FINE FRAGRANCES

Anyone watching the creation of a fine fragrance by a perfumer can't help but notice that there are a lot of experiments going on. Perfumers blend accords or keys, smell them (usually on specially created perfumery blotters), wait a while, smell again and move on. Occasionally, the perfumer will pause to write some notes, perhaps to highlight some interesting "notes" or smells that emerge from the blend. All in all, a lot of empirical experimentation goes on, most of which is not particularly systematic, but experimentation nonetheless.

Of course and as in every profession, skill and native ability play large roles. The perfumer is the corporation's secret weapon, the person who can create the pleasing smell that gives the product identity which may be a strong message and a sensory reinforcer. It is no wonder, therefore, that perfumers don't necessarily welcome experimental design, despite the fact that the perfumer learns by experimentation. Experimental design removes some of the mystique. Not a lot, mind you. The perfumer has to be good just to "get into the game". All that the RDE effort does is systematized. But we get ahead of the story.

Our third and final story takes place in the late days of 1978, at Chesebrough Pond's. More than 30 years ago, Chesebrough Pond's, now part of Unilever, was itself a mini-conglomerate comprising Pond's (well known for Pond's® Cream and for Vaseline® Intensive Care Lotion), as well as Aviance® fragrance and of course Ragu® Spaghetti Sauce. Quite a collection, of course, merged under the careful guidance of Ralph Ward, who would later sell this mini-conglomerate to Unilever.

During those fast-developing days of 1978–1979, when the food industry was starting to recognize the value of RDE and the cosmetic industry was quickly waking up to it, the fragrance industry was still fast asleep. Business in the fragrance industry was not done as it is today, on the basis of competitive submissions that are rigorously tested by consumers. Of course, there were competitive submissions from fragrance houses, but a lot of the business was done on the basis of relationships, often smoothed over by a three-course lunch that could last several hours and be completed as the afternoon wore on. The notion of RDE to create fragrances was not particularly welcome in such a world and indeed many of the perfumers recoiled in open horror (perhaps feigned) when the possibility of experimentation was even mentioned.

At that time, though, R&D was waking up. One of the senior marketers at Chesebrough, Joe Melnick, responsible for international marketing especially in South America, suggested to a fragrance and flavor supplier that it ought to consider quite seriously the potential for systematic variations in the fragrance submission. Melnick's reason was eminently simple: experimental design of fragrances could open the possibility to better understand fragrances, the reason for fragrance acceptance and the reduction of fragrance cost by the adjustment of the components, or "accords". To Melnick, it seemed perfectly reasonable to apply science to creativity and make both better.

Melnick's efforts led to one of the first, if not the actual first, experimental design of fragrances. Working with the technical staff of a major flavor and fragrance company, Melnick convinced the supplier company to do the experiment. The chief perfumer of the fragrance supplier company developed four different accords, blended them by the experimental design and the "game was afoot". The experiment itself was unremarkable, the results were crisp and the outcome was a better fragrance for Chesebrough Pond's.

Those were the initial results. More interesting for our story were the repercussions and then the long-term changes. During a meeting of the Society for Cosmetic Chemists, somehow the word got out and appeared in one of the trade newsletters: there was a new approach of "fragrance by numbers". The reaction initially was of disbelief, then horror. The disquiet and hubbub soon quieted down and was forgotten. But, it wasn't forgotten by the fragrance company and its competitors, who began finally to use experimental design and consumer research as a tool for development. What had been rejected initially, what had seemed a source of irritated astonishment, turned out to be one of the more important tools for the field. But of course, no one wanted to talk about that.

DISCUSSION AND CONCLUSIONS

In this very short history of how RDE came to be, the focus is on history, on the stories of the science and method, not on the method itself. One should always keep in mind that in the development of scientific methods and world views, not all is formal, well-ordered, hypothetico-deductive reasoning and well-executed, but soulless, experiments. Reading the literature of a field, one almost gets the sense in many fields that it was invented automatically, with no problems, no *sturm und drang*, aggravations, arguments, or passions.

The RDE did not appear like Athena fully formed from the head of Zeus. It may seem from reading the scientific and technical literature that RDE developed gradually, as a response both to the scientific development of the times and as a gentle response to ongoing business issues that pleasantly presented themselves, asking for a solution. Nothing could be further from the truth. The actual history of RDE, if there may be such as a thing as actual history, is more a story of conflicts, resolutions, statistical fights, corporate needs, internecine warfare and finally the acceptance of the future and disciplined research simply as inevitable.

Wars are often recounted by the later generations who were not involved, who did not suffer. To these later generations, the story of the war is just that—a story—fact after fact, compiled in a rigorous manner to create a history. And so, perhaps will the story of RDE eventually become that history. Yet, it is important to bear in mind as you read the science and methods contained in this edited book that you are reading a history in the making, a discipline that has great promise, a world view that may come to dominate thinking for years to come as it spreads its wings from basic science, statistics, perfumes and sweeteners, to the body politic and social policy.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Box, G. E. P., Hunter, J., & Hunter, S. (1978). *Statistics for experimenters*. New York: Wiley and Sons, Inc.

Cloninger, M. & Baldwin, R. (1970). Aspartylphenylalanine methyl ester: A low-calorie sweetener. *Science, 170*(3953), 81–82.

Hootman, R. (1992). *Manual on descriptive analysis testing for sensory evaluation.* Philadelphia, PA: ASTM.

Jacoby, J. (1978). Consumer research: A state of the art review. *The Journal of Marketing, 42*(2), 87–96.

Moskowitz, H. R. (1983). *Product testing and sensory evaluation of foods: Marketing and R&D approaches*. Westport, CT: Food & Nutrition Press.

Moskowitz, H. R. (1984). *Cosmetic product testing: A modern psychophysical approach*. New York: Marcel Dekker, Inc.

Moskowitz, H. R. (1985). *New directions in product testing and sensory evaluation of foods*. Westport, CT: Food & Nutrition Press.

Moskowitz, H. R. (1994). *Food concepts and products: Just in time development*. Trumbull, CT: Food and Nutrition Press.

Moskowitz, H. R. (1996). *Consumer evaluation and testing of personal care products*. New York: Marcel Dekker.

Moskowitz, H., Wolfe, K., & Beck, C. (1979). Sweetness and acceptance optimization in cola flavored beverages using combinations of artificial sweeteners—a psychophysical approach. *Journal of Food Quality, 2*(1), 17–26.

Plackett, R. & Burman, J. (1946). The design of optimum multifactorial experiments. *Biometrika, 33*, 305–325.

Rao, R. (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *Supplement to the Journal of the Royal Statistical Society, 9*(1), 128–139.

Rockett, F. (1956). Speed of form recognition as a function of stimulus factors and test anxiety. *The Journal of Abnormal and Social Psychology, 53*(2), 197–202.

Stevens, S. S. (1975). *Psychophysics: An introduction to its perceptual, neural and social prospects*. New York, NY: Wiley and Sons.

Stone, H. & Sidel, J. (2004). *Sensory evaluation practices* (3rd ed.). Maryland Heights, MO: Academic Press.

Yates, F. (1964). Sir Ronald Fisher and the design of experiments. *Biometrics, 20*(2), 307–321.

CHAPTER 9

Sensory Optimization in Research and Development

Maximo Gacula Jr.*

Gacula Associates Consulting, Scottsdale, Arizona, USA

Abstract: The advantages and pitfalls of sensory optimization are outlined in this paper. The response surface design is presented in terms of application to optimization studies. These studies are done by research and development. They deal with issues such as the role of ingredients in a product formulation and the discovery of optimal combinations of these ingredients that generate the desired sensory properties. Contour maps developed from the experiment illustrate how to discover the best ingredient combinations and how to avoid extrapolating beyond the range tested in the actual experiment.

Keywords: Optimization technique, contour map, response surface design, mixture design, normal probability plot.

INTRODUCTION

There are two main issues in the application of sensory optimization techniques in research and development (R&D) that need to be addressed and resolved. The first issue is the researcher's tendency to remain with the traditional method of "one experiment at a time". The second issue how to communicate to R&D that the bottom-line purpose of design of experiments is cost reduction. Although old, these issues are still encountered in practice. Yet several companies have eliminated these issues in R&D and manufacturing. With the tremendous advancement of computer technology, resolution of these issues can be facilitated more broadly. The subject of this chapter is how to extend the value of experimentation across companies in a palatable fashion.

USEFULNESS OF OPTIMIZATION TECHNIQUES

In general, optimization in research comprises a series of steps to obtain the best result at least cost under a given set of circumstances. The main component of

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

optimization is design of experiments, where various designs exist corresponding to the purpose of the study. The statistical aspects of experimental design can be found in various applied books (Montgomery, 1991; Myers and Montgomery, 1995). Considerations in the world of sensory evaluation are available (Gacula, 1993; Gacula and Singh, 1984; Gacula, Singh and Altan, 2009). The advantages and pitfalls of optimization techniques were discussed in Gacula (1993) and are given below with some expansions.

Advantages:

1. The optimization method enforces discipline in the conduct of scientific research, from conception of the project to execution and through analysis. Most importantly, optimization generally yields quality data.

2. Optimization is fast and cost-effective by avoiding experimental re-runs. Re-runs occur when the method of "one experiment at a time" is adopted for a particular project.

3. Optimization is statistically efficient because the data are modeled by equations that summarize the relations in the data. In the regression model, depending on the design, it includes the linear effect, quadratic effect, interaction effect and error (lack of fit and experimental error).

4. Optimization generates and provides knowledge, a database that one can use in order to answer direct "what if" questions such as: What if ingredient or process X becomes expensive and it is desirable to reduce the amount in the formulation—what will then happen to the sensory and/or physical characteristics of the product? It is not necessary to run a full-pledge experiment as the database can provide the necessary information. Conducting a validation run is sufficient.

5. Optimization results often uncover several potential product formulas for consumer evaluation.

6. Optimization results provide direction to R&D to meet the changing market demands. The needed information can be obtained in the response surface map.

However, there are pitfalls in the use of the optimization method when the researcher is not careful during the planning stage of the study. It is important that a team must be formed to define the objectives of the study. Sometimes the following problems are encountered, usually because not enough care has been taken at the start of the project to understand the problem in sufficient detail:

1. Important factors/variables that affect the response being measured have not been correctly identified. The investigator should know the function of each ingredient in the formulation or have some hypothesis regarding their effects on the response.

2. The lower and upper levels of the factors (ingredients) have been incorrectly specified. Preliminary work is needed when these levels are not available. Some estimates may be obtained from existing products and research experience. On a practical note, such estimates should be inserted as the middle level in the design specification.

3. Overuse of extrapolation of response surface or contour maps without checkpoints. This occurs when there is more than one optimal point or area in the response surface.

4. Use of incorrect statistical models and experimental designs.

5. Failure to verify the correctness of the selected optimum formulas against a control or "gold standard" in the marketplace. The verification should be done before the final decision on product plant production.

RESPONSE SURFACE DESIGN

We now move to the response surface method (RSM). In this method, the factors or variables under the experimenter's control are unrestricted and independent of each other. *That is, changing the in-going levels of one variable does not affect the levels of the other variables in the formulation*.

An example of RSM can be seen when working with a product with 10 ingredients in its formulation. One may wish to change the amount of the three ingredients as follows:

Ingredient	Amount (lb)
A	1.0
B	0.5
C	3.0
Total	4.5

It is straightforward to change the amount of any ingredient (A, B, or C) without affecting the two variables. The only value affected is the total amount of the formulation, which may either increase or decrease, depending on whether one is adding or subtracting an amount. The experimental designs commonly used for nonmixture experiments are the Box–Wilson, popularly known as the central composite design; Plackett–Burman; and the Box–Behnken designs. However, in general, all factorial designs can be used in response surface experiments. Response surface experiments are common when optimizing processing variables, *i.e.* time and temperature. In food formulation, the amount of nitrate (ppm) and nitrite (ppm) in the formulation can be varied to control microbial activities.

One off-the-shelf program, Scantron® Design-Expert (Stat-Ease Inc., 1997) gives a lot of design choices as will be illustrated in the examples to follow.

EXAMPLE 1

In this example, there are 10 ingredients in the formula already in the market. Because of quality problem and production cost when compared with a competitor, it was decided to look at three ingredients: X, Y and Z. The prescribed ingredient limits are as follows (Table **1**).

Table 1: Levels of Three Ingredients (X,Y,Z) Studied in an RSM Experiment.

Ingredients	**Low %**	**Middle %**	**High %**
IngX	0.0	1.5	3.0
IngY	10.0	18.0	26.0
IngZ	0.0	2.0	4.0

The middle level is not a required input of Design-Expert. But for information purposes for the R&D scientist, it is recommended that this value be tested. In most cases, the middle level is the existing level of the current product. Being a response surface or a nonmixture experiment, the total amount (%) of X, Y and Z in the formulation is not specified.

The next step is to access Design-Expert. For a “small” central composite design as the choice, 15 formulations or design points were generated, containing five center points (formulations 11–15; Table **2**). For cost reasons and without sacrificing the quality of the data, three of the five center points can be deleted to generate 12 formulations to be produced in the laboratory. In cases where the lower limit is zero, negative value can occur in the design, generated by the design program. The negative value can be brought to zero without serious effect on the design. Likewise, the prescribed upper limit can be exceeded in the generated design and no changes should be made. These situations can be seen in Table **2**.

Table 2: Design Generated by Design-Expert.

Formulation	Random	Block	IngX	IngY	IngZ
1	15	Block 1	3.00	26.00	0.00
2	1	Block 1	3.00	10.00	4.00
3	3	Block 1	0.00	26.00	4.00
4	10	Block 1	0.00	10.00	0.00
5	7	Block 1	0.62	18.00	2.00
6	6	Block 1	3.62	18.00	2.00
7	5	Block 1	1.50	6.69	2.00
8	14	Block 1	1.50	29.31	2.00
9	11	Block 1	1.50	18.00	0.83
10	9	Block 1	1.50	18.00	4.83
11	4	Block 1	1.50	18.00	2.00
12	12	Block 1	1.50	18.00	2.00
13	8	Block 1	1.50	18.00	2.00
14	13	Block 1	1.50	18.00	2.00
15	2	Block 1	1.50	18.00	2.00

Fig. **1** provides a view of the 15 formulations in the design map for IngY and IngX. Formulations 9 and 10 happen to be in the center point as IngZ is not shown on the map. However, Table **1** shows that in fact formulations 9 and 10 differ in IngZ.

The importance of the design points (formulations) away from the center provides various directions of the effects in the model for optimizing the best combinations of the three ingredients. As indicated earlier, the critical choice of prescribing the lower and upper limits in the design is apparent. The SAS© program "code" to produce the plot in Fig. **1** is given in Table **3**. The user is encouraged to go over the program code and program statements (SAS, 1999).

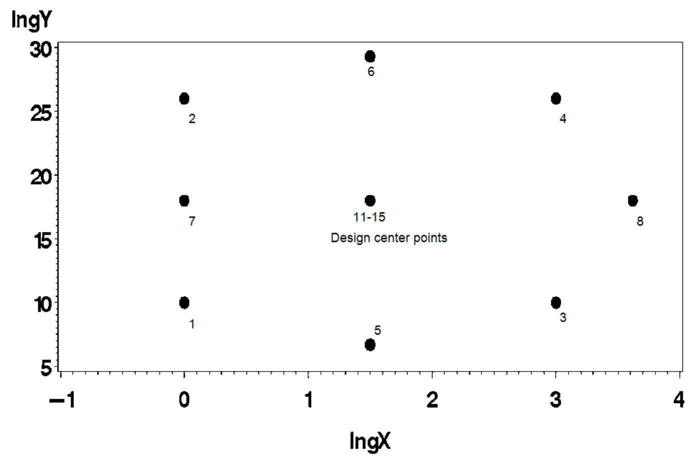


Figure 1: Design map for a central composite design showing only two ingredients for simplicity.

Table 3: SAS© Program Code for Obtaining the Formulation Map.

```
*PROG DOE PLOT.SAS;
OPTIONS NODATE;
DATA A;
INPUT FORMULATION INGY INGX;
CARDS;
1 10.00 0.00
2 26.00 0.00
3 10.00 3.00
4 26.00 3.00
5 6.69 1.50
6 29.31 1.50
7 18.00 0.00
8 18.00 3.62
9 18.00 1.50
10 18.00 1.50;
GOPTIONS RESET=GLOBAL GUNIT=PCT
FTEXT=SWISSB HTITLE=3 HTEXT=5;
SYMBOL1 COLOR=RED VALUE=DOT HEIGHT=5;
PROC GPLOT DATA=A;
PLOT INGY*INGX /
VAXIS=5 TO 30 BY 5
HAXIS=-1 TO 4 BY 1;
RUN;
```

Building the equation is straightforward with today's computational power and with canned programs available in statistical packages. The key decision to make is the form of the equation. Most research data follow the quadratic model. For our three-variable experiment we write the model as follows:

$$Y = B_o + B_1X + B_2Y + B_3Z + B_{11}X^2 + B_{22}Y^2 + B_{33}Z^2 + B_{11}XY + B_{22}XZ + B_{33}YZ +$$

Random error,

where Y = response being measured in the experiment,

B_o = intercept,

$B_1X + B_2Y + B_3Z$ = linear effects,

$B_{11}X^2 + B_{22}Y^2 + B_{33}Z^2$ = quadratic effects,

$B_{11}XY + B_{22}XZ + B_{33}YZ$ = interaction effects.

Sometimes one or more of the effects is not statistically significant. When the term is not significant, it may be deleted from the equation, but does not have to be.

Table **4** shows the sensory data for this example. In this example, formulations 11 and 12 are center points in the design as the other three points were deleted. A higher texture score is desirable in this scale.

Table 4: Texture Score (7-Point Scale) of Each Formulation with Two Center Points.

Formulation	Texture Score
1	4.2
2	3.5
3	3.1
4	6.3
5	3.0
6	5.5
7	3.0
8	4.9
9	4.2

Table 4: cont….

10	4.5
11*	5.0
12*	5.2

Note. *Center point.

The Design-Expert program is interactive. Many of its outputs can be printed. When trained to use this software, the analyst can obtain a great deal of useful information from the study, often leading to a profound knowledge of the product. We will confine our presentation of data to a limited amount of output, as this chapter is one on design, rather than a tutorial on reading output.

CREATING A PRODUCT MODEL

The primary output of the software is the statistical model and optimization plots. The outputs provide two important pieces of information that the analyst most needs to know: the role of the ingredients and the optimal combinations of these ingredients to satisfy the texture requirement of the product. Table **5** shows the analysis of variance and the statistical model used to generate the contour plots. Design-Expert includes important statistical explanations in the output and other steps that can be done as shown in this table. Like many integrated programs, the design is intimately linked with the subsequent analysis. Hence for this experiment, the design, *i.e.* layout of the combinations, should be generated by the Design-Expert before the data can be analyzed.

Table 5: Output from Design-Expert

Response: Texture

ANOVA for Response Surface 2FI Model

Analysis of variance table [Partial sum of squares]

Source	Sum of Squares	DF	Mean Square	F Value	Prob > F
Model	10.89	6	1.82	5.42	0.0419 significant
A	2.96	1	2.96	8.83	0.0311
B	1.81	1	1.81	5.39	0.0680
C	0.046	1	0.046	0.14	0.7270

Table 5: cont….

AB	2.06	1	2.06	6.16	0.0557
AC	3.36	1	3.36	10.04	0.0249
BC	3.45	1	3.45	10.29	0.0238
Residual	1.68	5	0.34		
Lack of Fit	1.66	4	0.41	20.69	0.1632 not significant
Pure Error	0.020	1	0.02		
Cor Total	12.57	11			

The Model F-value of 5.42 implies the model is significant. There is only a 4.19% chance that a "Model F-Value" this large could occur due to noise.

Values of "Prob > F" less than 0.0500 indicate model terms are significant.

In this case A, AC, BC are significant model terms.

Values greater than 0.1000 indicate the model terms are not significant.

If there are many insignificant model terms (not counting those required to support hierarchy), model reduction may improve your model.

The "Lack of Fit F-value" of 20.69 implies the Lack of Fit is not significant relative to the pure error. There is a 16.32% chance that a "Lack of Fit F-value" this large could occur due to noise. Non-significant lack of fit is good—we want the model to fit.

Std. Dev.	0.58	R-Squared	0.8667
Mean	4.37	Adj R-Squared	0.7067
C.V.	13.26	Pred R-Squared	0.6359
PRESS	4.58	Adeq Precision	7.238

The "Pred R-Squared" of 0.6359 is in reasonable agreement with the "Adj R-Squared" of 0.7067.

“Adeq Precision” measures the signal to noise ratio. A ratio greater than 4 is desirable. Your ratio of 7.238 indicates an adequate signal. This model can be used to navigate the design space.

Coefficient Factor Estimate	**DF**	**Standard Error**	**95% CI Low**	**95% CI High**	**VIF**
Intercept 4.33 1	0.17	3.90	4.76		
A-IngX 1.00	1	0.33	0.13	1.86	2.34
B-IngY 0.67	1	0.29	-0.072	1.42	2.00
C-IngZ 0.12	1	0.33	-0.74	0.98	2.34
AB 1.10	1	0.44	-0.039	2.24	2.34
AC 1.30 1	0.41	0.24	2.35	2.00	
BC 1.42 1	0.44	0.28	2.56	2.34	

Final Equation in Terms of Coded Factors:

Texture =
+4.33

+1.00 * A
+0.67 * B
+0.12 * C
+1.10 * A * B
+1.30 * A * C
+1.42 * B * C

Final Equation in Terms of Actual Factors:

Texture =
+8.66197
-1.84918 * IngX
-0.23089 * IngY
-2.18417 * IngZ
+0.091562* IngX * IngY
+0.43225* IngX * IngZ
+0.088759* IngY * IngZ

Diagnostics Case Statistics

Standard Order	Actual Value	Predicted Value	Residual	Leverage	Student Residual	Cook's Distance	Outlier t
1	4.20	4.25	-0.053	0.834	-0.225	0.036	-0.202
2	3.50	3.55	-0.053	0.834	-0.225	0.036	-0.202
3	3.10	3.15	-0.053	0.834	-0.225	0.036	-0.202
4	6.30	6.35	-0.053	0.834	-0.225	0.036	-0.202
5	3.00	3.33	-0.33	0.442	-0.770	0.067	-0.734
6	5.50	5.74	-0.24	0.721	-0.770	0.219	-0.734
7	3.00	3.38	-0.38	0.584	-1.013	0.206	-1.016
8	4.90	5.28	-0.38	0.584	-1.013	0.206	-1.016
9	4.20	4.20	-4.303E-003	0.442	-0.010	0.000	-0.009
10	4.50	4.50	-3.043E-003	0.721	-0.010	0.000	-0.009
11	5.00	4.33	0.67	0.084	1.213	0.019	1.292
12	5.20	4.33	0.87	0.084	1.574	0.033	1.982

Proceed to Diagnostic Plots (the next icon in progression). Be sure to look at the:

1) Normal probability plot of the studentized residuals to check for normality of residuals.
2) Studentized residuals *versus* predicted values to check for constant error.
3) Outlier t *versus* run order to look for outliers, *i.e.* influential values.
4) Box-Cox plot for power transformations.

If all the model statistics and diagnostic plots are OK, finish up with the Model Graphs icon.

Let's look at the output of the model in a bit of detail to get a sense of what emerges from the analysis:

1. As shown in the analysis of variance, the effect of IngX denoted by A on texture is significant ($P < 0.0311$), effect of IngY(B) is directional ($P < 0.0680$) and IngZ(C) has insignificant effect.

2. The interactions of these three ingredients are directional and significant, making the contour plot very informative. The contour plot or map shows these interactions on texture to be clearly at various combinations of the ingredient levels.

3. The important information in the analysis is the normal plot of residuals (Fig. **2**), which provides the goodness of fit of the model that was estimated by statistical procedures to describe how the responses co-vary with the ingredients in the 12 formulations. As shown in Fig. **2**, the fit is satisfactory, with the formulation residuals appearing to form a straight line.

The next important outputs are the contour maps. These contour maps show the combination of two variables generating a constant response. We follow these steps:

1. In our study we have three variables. We first select the magnitude of the dependent variable. That magnitude remains constant. In Fig. **3a** and **3b**, we see numbers atop a contour. Each number corresponds to the one specific magnitude of the dependent variable. That magnitude will be constant for a single contour. Note also that each figure comprises a series of contours, all wrapping around each other.

2. We deal with three variables, but our plot only shows two variables. Consequently, we hold one of the variables constant.

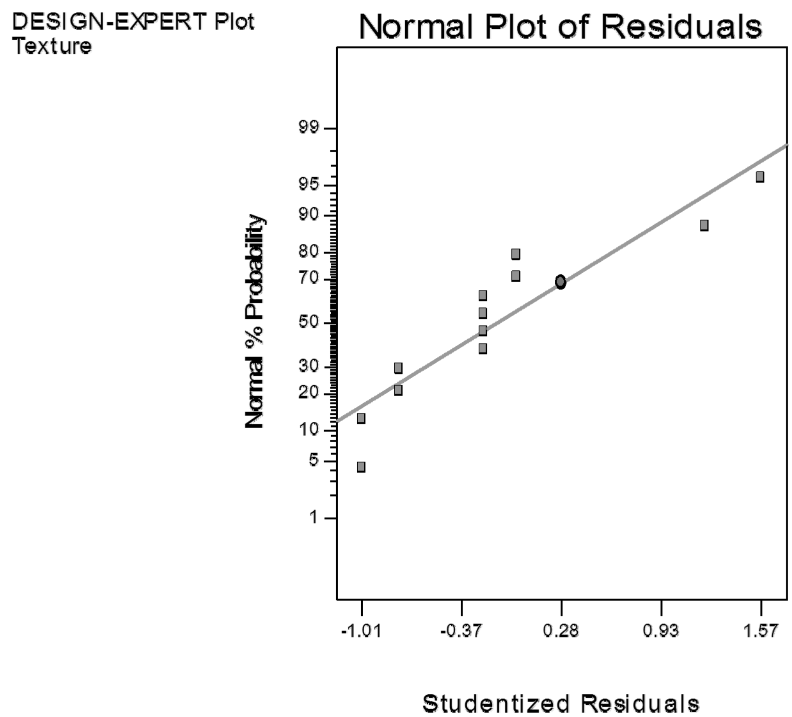


Figure 2: Residual plot of the 12 formulation design points.

3. We then systematically change one of the two remaining variables, from low to high in very small increments.

4. For each change in the variable in Step 3 above, we know the level of the variable that we just changed, the level of the variable held constant and of course the level of the dependent variables (*i.e.* the number on the contour). That information suffices to estimate the level of the remaining second variable.

5. We now estimate what the second remaining variable would have to be to generate the desired response at various combinations of the ingredients.

6. The contour lines were obtained by the statistical model given in Table **6**.

Table 6: Optimization Solutions Where the Texture Score Was Set at 5.0–6.5 Range as the Target Response

Constraints		**Lower Limit**	**Upper Limit**	**Lower Weight**	**Upper Weight**
Name	**Goal**				
IngX	is in range	0	3	1	1
IngY	is in range	10	26	1	1
IngZ	is in range	0	4	1	1
Texture	maximize	5	6.5	1	1
Solutions Number	**IngX**	**IngY**	**IngZ**	**Texture**	**Desirability**
1	2.86	25.68	1.99	6.8	1.000
2	2.74	24.38	3.36	8.0	1.000
3	2.85	19.50	3.66	6.8	1.000
4	2.68	22.26	3.24	7.1	1.000
5	2.23	25.86	2.70	6.7	1.000
6	2.81	22.27	3.92	8.0	1.000
7	0.00	10.00	0.00	6.4	0.902
7 Solutions found					

That equation is written as follows:

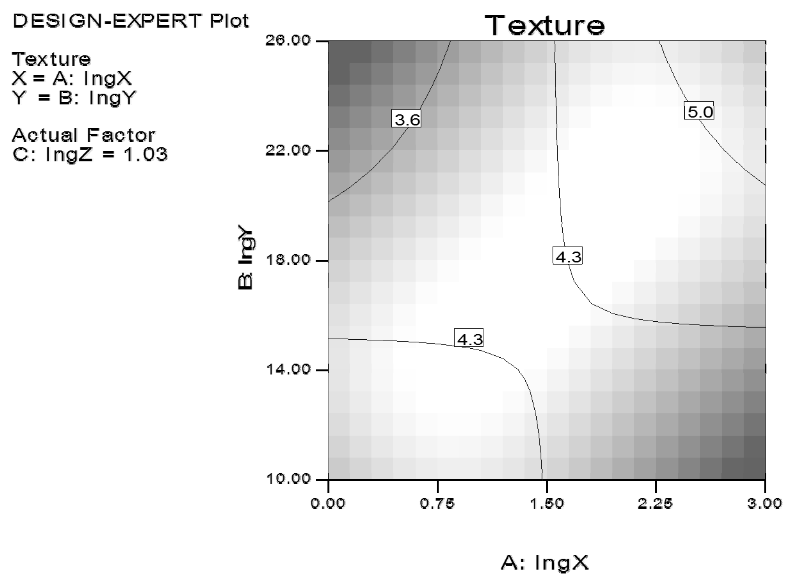
$$\mathbf{Texture = 8.662 - 1.849(IngX) - 0.231(IngY) - 2.184(IngZ) + 0.092(IngX)(IngY) + 0.432(IngX)(IngZ) + 0.089(IngY)(IngZ)}$$

7. Notice that the model contains only linear effects and interaction effects. There are no square, *i.e.* quadratic, effects. Apparently, the quadratic effect was not important and was automatically excluded by Design-Expert.

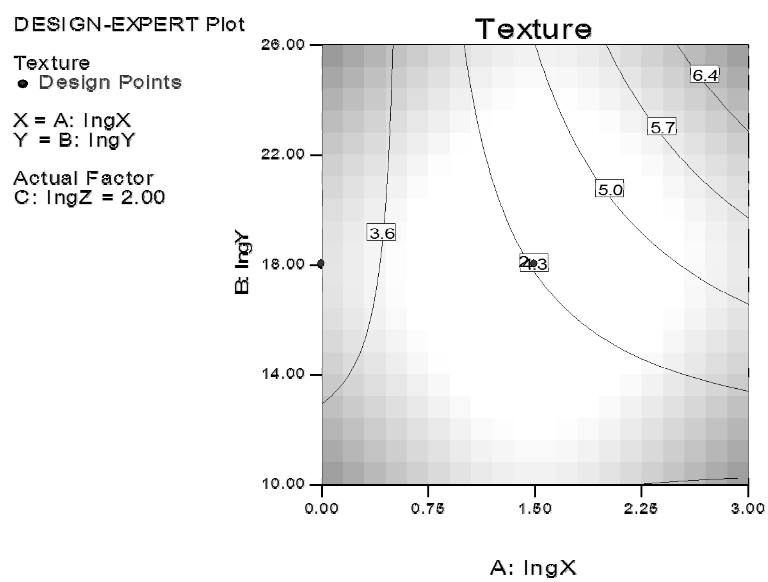
8. We can now systematically explore the results by making different assumptions. The two key assumptions are the level of IngZ (which we set as constant) and the magnitude of the rating (which we also set as constant).

9. By setting IngZ = 1.03, we find the contour lines that show the area containing undesirable texture scores (Fig. **3a**). By setting IngZ = 2.00, we find that the contour lines with desirable scores (above 5.0) start to appear (Fig. **3b**). By setting IngZ = 4, we find the region that clearly stabilizes the location of the desirable area (Fig. **3c**).

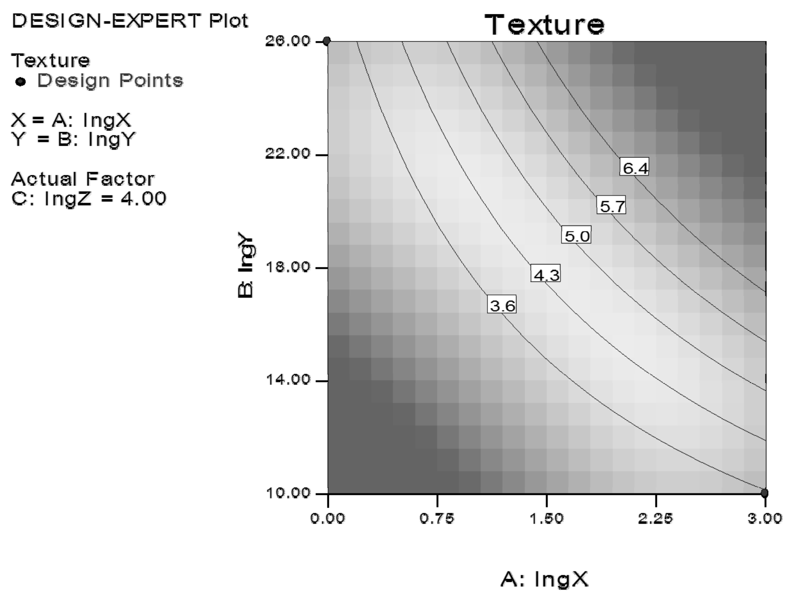
10. Our conclusion from the exploration of the response surface through these contour maps is that we find several combinations of ingredient levels that generate optimal formulations.



(a)



(b)



(c)

Figure 3: (a) Contour map with ingredient IngZ = 1.03 (between lower limit and middle level). (b) Contour map with ingredient IngZ = 2.00 (middle). (c) Contour map with ingredient IngZ = 4.00 (upper limit).

OPTIMIZING A FORMULATION

The next step is the optimization process, which gives the optimal solutions in numerical form (Table **6**). There are options that the analyst can set in terms of the range of the target score. On a 7-point scale, the logical range is 5.0–6.5. One can use the highest score of 7, but predicted and meaningless scores often occur above the scale limit.

The results in Table **6** show six workable optimal formulas. The seventh is excluded, because its value is less than 1.00. The next step is for the research analyst to calculate the cost of producing each formulation solution and select two to three least-cost formulations for confirmation testing. For example, if IngY is an expensive ingredient, we may choose solution 3 with 19.50%, resulting in 2.85% IngX and 3.66% for IngZ. These solutions generate a predicted texture score of 6.8.

The optimal solutions can be presented graphically. The advantage of this is that it also provides the area that one should avoid because it does not meet the texture score criterion. That criterion is that the solution must lie between a low of 5.0 and a high of 6.5.

Here is the way to interpret the results:

1. At IngZ = 1.03, map area below texture = 5.0 should be avoided (Fig. **4a**).

2. Increasing IngZ amount to 2.00, the desired area of the map becomes visible and a "flag" is indicated by following the Design-Expert instructions after a "right click" (Fig. **4b**).

3. A texture score of 6.5 can be achieved by the ingredient combination of (X, Y, Z) = (2.84%, 24.24%, 2.00%) as indicated in the "flag".

4. Setting the amount of IngZ to its maximum of 4.00%, the contour plot completes the location of the desirable area of the map (Fig. **4c**).

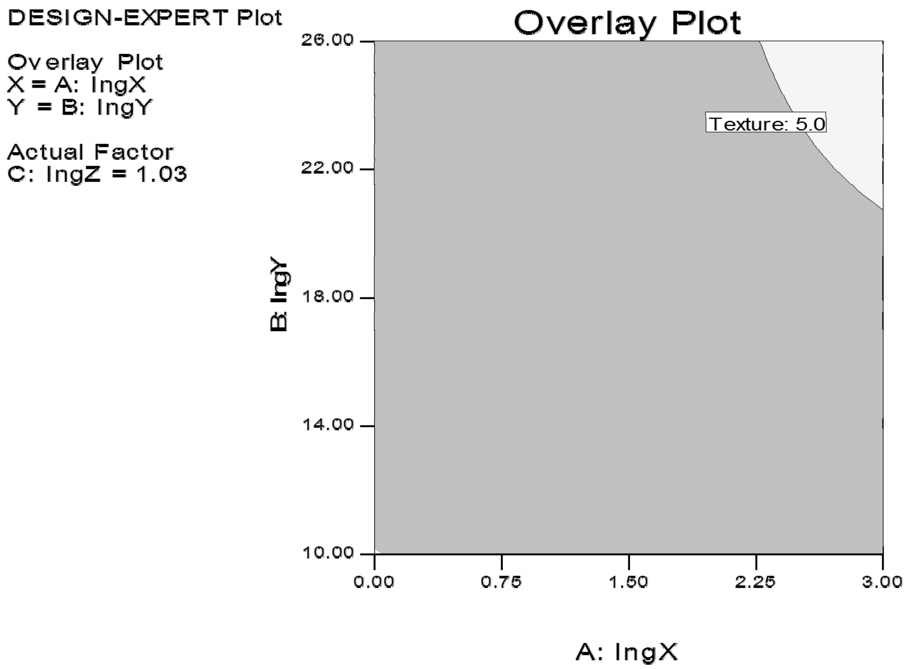
5. For information purposes, a "flag" was set at the bottom of the plot that results in texture score of 2.76 with (X, Y, Z) = (1.47%, 11.64%, 4.00%), which is obviously an undesirable ingredient combination.

6. In this example, one product optimization study gives three pieces of information that are critical to developers: (a) optimal formulations; (b) undesirable area in the map as defined by the target response; and (c) the specific least-cost formulations. These three pieces of information cannot be obtained by the one-experiment-at-a-time practice because they require that the three variables interact together, dynamically, with each variable independently moving in the appropriate range but all simultaneously contributing to the measured response(s).

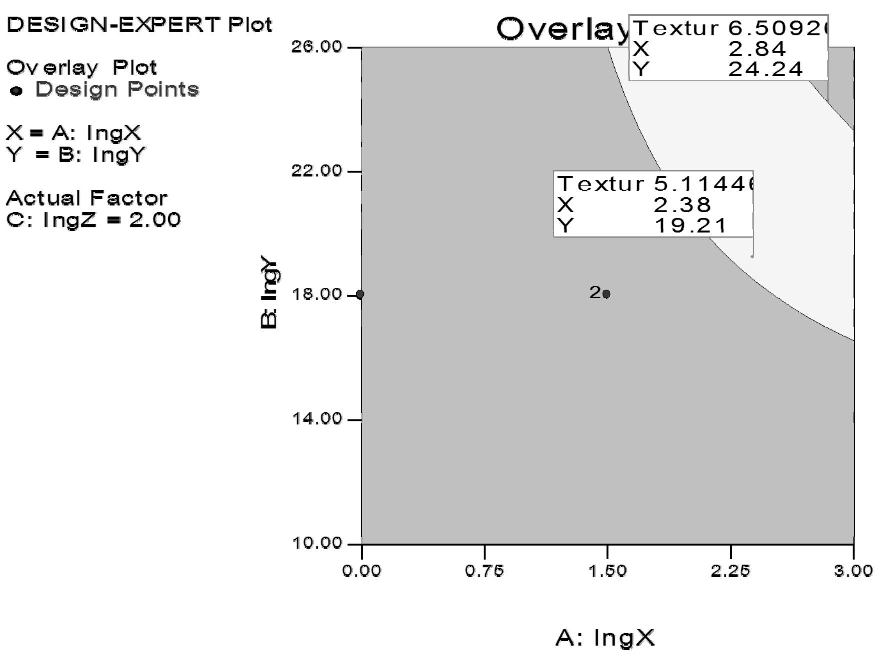
EXAMPLE 2

Historical Data in Place of Experimentally Designed Studies

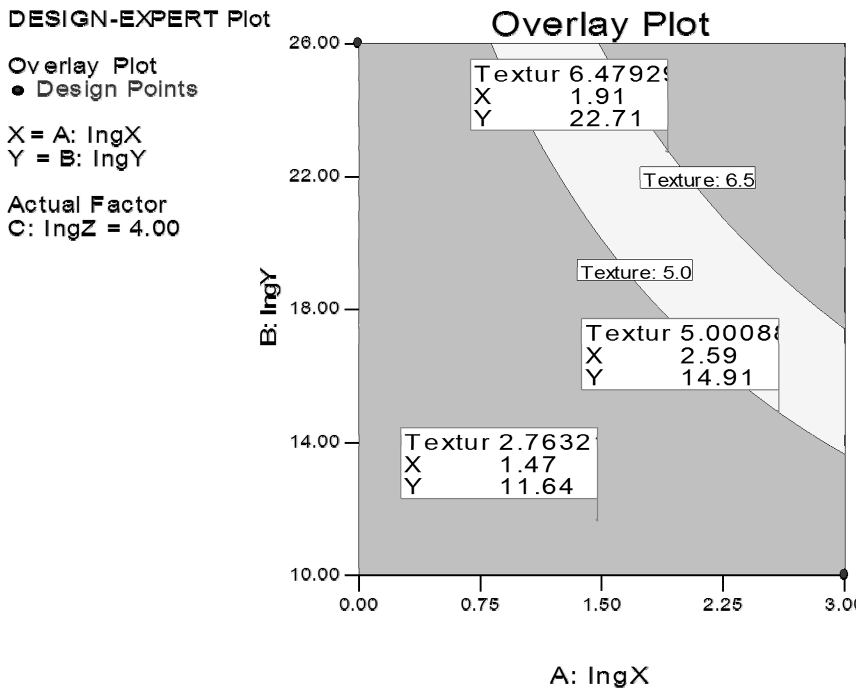
This example illustrates optimization of historical data generally found in R&D files. Optimization of these data cannot be done using Design-Expert. An SAS code is used instead.


DESIGN-EXPERT Plot
Overlay Plot
X = A: IngX
Y = B: IngY
Actual Factor
C: IngZ = 1.03
Overlay Plot
Texture: 5.0
26.00
22.00
18.00
14.00
10.00
B: IngY
0.00
0.75
1.50
2.25
3.00
A: IngX

(a)


DESIGN-EXPERT Plot
Overlay Plot
Design Points
X = A: IngX
Y = B: IngY
Actual Factor
C: IngZ = 2.00
Overlay
Textur 6.5092
X 2.84
Y 24.24
Textur 5.1144
X 2.38
Y 19.21
2
26.00
22.00
18.00
14.00
10.00
B: IngY
0.00
0.75
1.50
2.25
3.00
A: IngX

(b)



(c)

Figure 4: (a) Optimization contour map. (b) Optimization contour map. (c) Optimization contour map.

We begin with the historical data. Table **7** contains the historical data of eight formulas taken in the last 10 years. Notice that there is no experimental design in the data. There are variations, however, in the formulation.

Table 7: Overall Liking Mean Score for Each Formula and its Corresponding Ingredient Level (Ing A, Ing B, Ing C, Ing D).

Formulas	Ing A	Ing B	Ing C	Ing D	Overall Liking
1	2.07	2.93	15.00	0.0	5.33
2	2.34	3.66	13.70	2.0	6.33
3	1.19	5.81	13.50	0.0	5.73
4	1.19	5.81	13.70	2.0	5.80
5	3.50	1.50	10.90	4.0	5.53
6	1.90	5.10	9.40	4.0	5.80
7	2.61	4.39	15.00	0.0	6.13
8	3.50	1.50	13.70	2.0	6.07

It's now time to build a model. We don't need an experimental design to build the model. However, we have to make sure that the variables are reasonably independent of each other. ("Reasonably" here means that the variables are not correlated or multi-collinear).

In many of today's statistical programs we do not even have to create the equation. The program fits an equation to the data, or more accurately, fits a smoothing program to the data. The output is a surface map, of the same two-dimensional type we saw above, but where the ingredients were varied by experimental design. Here we wait for the computer program to create a smooth surface map. This map was done using the following ingredient combinations and holding two ingredient levels constant:

A *vs.* B holding C and D constant at various levels in each run

A *vs.* C holding B and D constant at various levels in each run

B *vs.* C holding A and D constant at various levels in each run

Table **8** shows the SAS code that creates the surface.

Table 8: SAS Code for A *vs.* B with C and D Values Interactively Changing at Each Run.

```
*PROG HISTORICAL DATA AB.SAS;
OPTIONS NODATE;
DATA RESPONSE;
INPUT FORMULA A B C D X1;
LABEL
A='ING A'
B='ING B'
C='ING C'
D='ING D'
X1='OVERALL LIKING'
;
CARDS;
1    2.071   2.929   15.00   0.0   5.33
2    2.343   3.657   13.70   2.0   6.33
3    1.186   5.814   13.50   0.0   5.73
4    1.186   5.814   13.70   2.0   5.80
5    3.500   1.500   10.90   4.0   5.53
6    1.900   5.100   9.40    4.0   5.80
7    2.614   4.386   15.00   0.0   6.13
8    3.500   1.500   13.70   2.0   6.07
;
```

Table 8: cont….

```
DATA GRID;
DO;
X1=.;
C=15;
D=4;
DO A=1 TO 4 BY.5;
DO B=1 TO 6 BY.5;
OUTPUT;
END;END;END;
DATA GRID;
SET RESPONSE GRID;
RUN;
PROC RSREG DATA=GRID OUT=PREDICT;
MODEL X1= A B C D / PREDICT;
RUN;
DATA PLOT;
SET PREDICT;
IF C=15;
IF D=4;
PROC G3D DATA=PLOT;
PLOT A*B=X1 / ROTATE=38
TILT=75 XTICKNUM=3 YTICKNUM=3
ZMAX=7 ZMIN=0 CTOP=RED CBOTTOM=BLUE CAXIS=BLACK;
RUN;
```

The SAS code in Table **8** contains some statements that are underlined. These are the statements that interactively change when the "plot data" change. These statements are linked to the "plot" statement in the "PROC G3D" procedure. The "zmax=7" indicates the high range of value of overall liking; "zmin=0" is the minimum value especially for rating scale that starts at 0. The procedure PROC RSREG outputs the regression parameters given in Table **8** for obtaining the predicted overall liking using SAS code in Table **9**:

Table 9: Abbreviated Output from Running the SAS Code Shown in Table **8.**

```
Parameter Estimate
Parameter DF Estimate from Coded Data
Intercept 1 -40.887231 6.025000
a 1 25.244980 0.776251
b 1 7.382097 0.641251
c 1 0.203333 0.569333
d 1 0.014667 0.029333
a*a 1 -3.517288 -4.708414
b*a 1 -2.212757 -5.522264
b*b 1 -0.259820 -1.208849
```

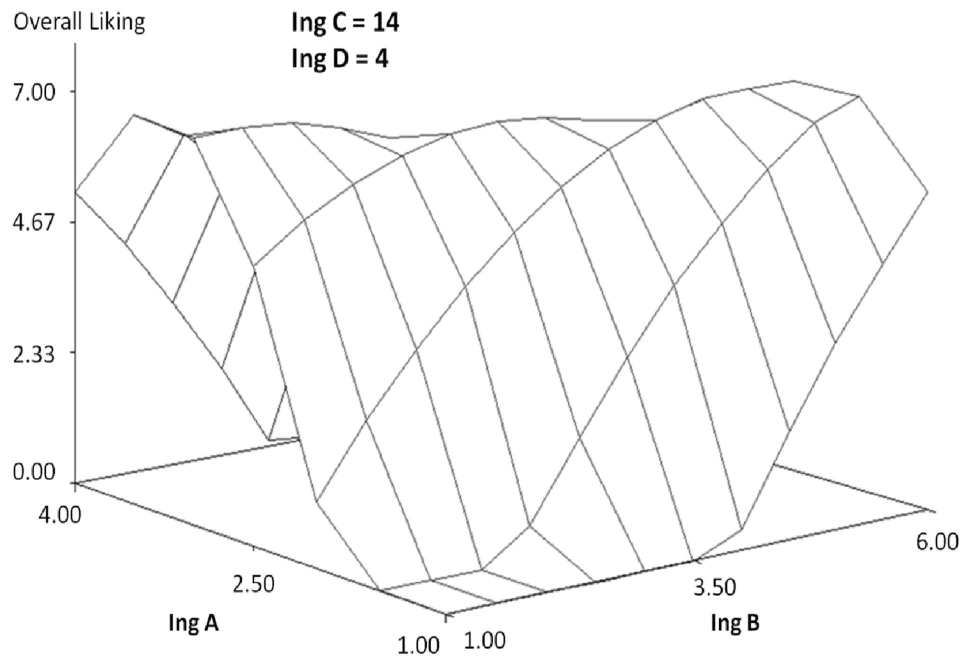
Note. The code generates the parameter estimates used to predict overall liking, based on the quadratic response surface regression model.

The response surface map emerging from the SAS codes in Tables **8** and **10** (tabulated part of the figure) appears in Fig. **5a**.

1. Another map can be obtained by changing the values for ingredients C and D in the SAS code.
2. The location of the predicted overall liking in the map can be manually identified from the three-dimensional map.
3. A similar manual estimate can also be done for Fig. **5b**, in particular Fig. **5c**.
4. As a check, the SAS code in Table **10** computes the predicted overall liking score.
5. The CARDS statement contains the data for obtaining the predicted score of interest.
6. The overall liking response surface regression consists of linear effects (a, b, c, d), quadratic (a*a, b*b) and cross product (a*b).
7. As shown in the map, the SAS terminology to describe the map is "Stationary point is a saddle point" (Fig. **5a**, **5b**) or "Stationary point is in a flatness" (some areas in Fig. **5a**, **5b**).

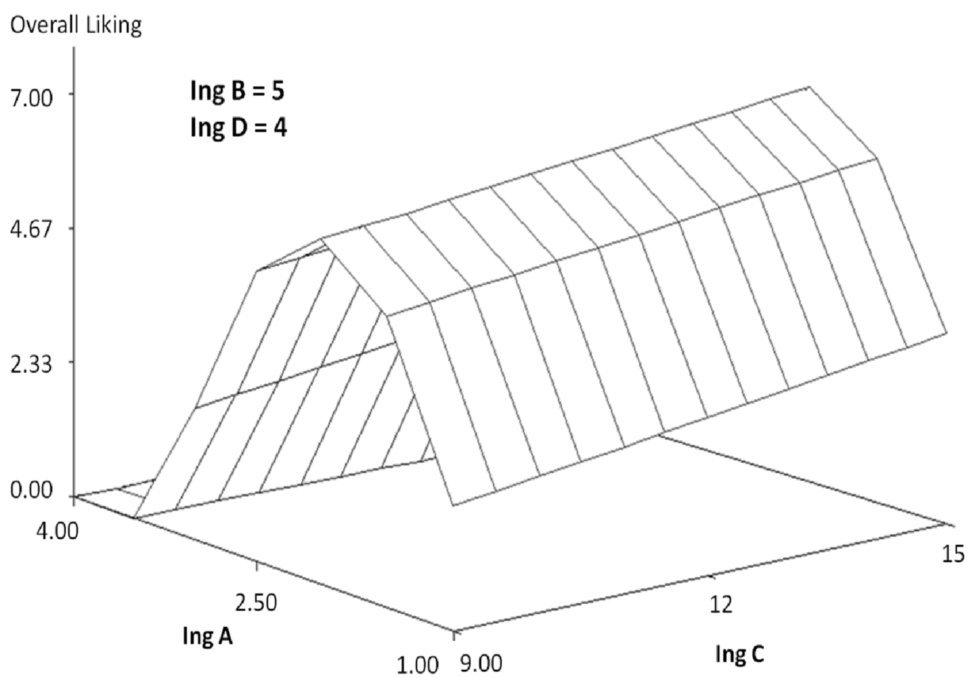
Table 10: SAS Code Containing Overall Liking Equations Obtained from the SAS Code Given in Table **8** to Produce the Response Surface Map in Fig. **5**.

```
*PROG PREDICTED VALUES AB.SAS;
OPTIONS NODATE;
DATA A;
INPUT A B C D;
S1=A*A;
S2=A*B;
S3=B*B;
OVERALLLIKING = -40.9 + A*25.2 + B*7.4 + C*0.203 + D*0.0145
+ S1*-3.52 + S2*-2.21 + S3*-0.260;
CARDS;
2.5 4.0 15 4
3.5 1.5 15 4
;
PROC PRINT DATA=A;
VAR A B C D OVERALLLIKING;
TITLE'PREDICTED VALUES A VS. B CONSTANT C AND D ';
RUN;
```



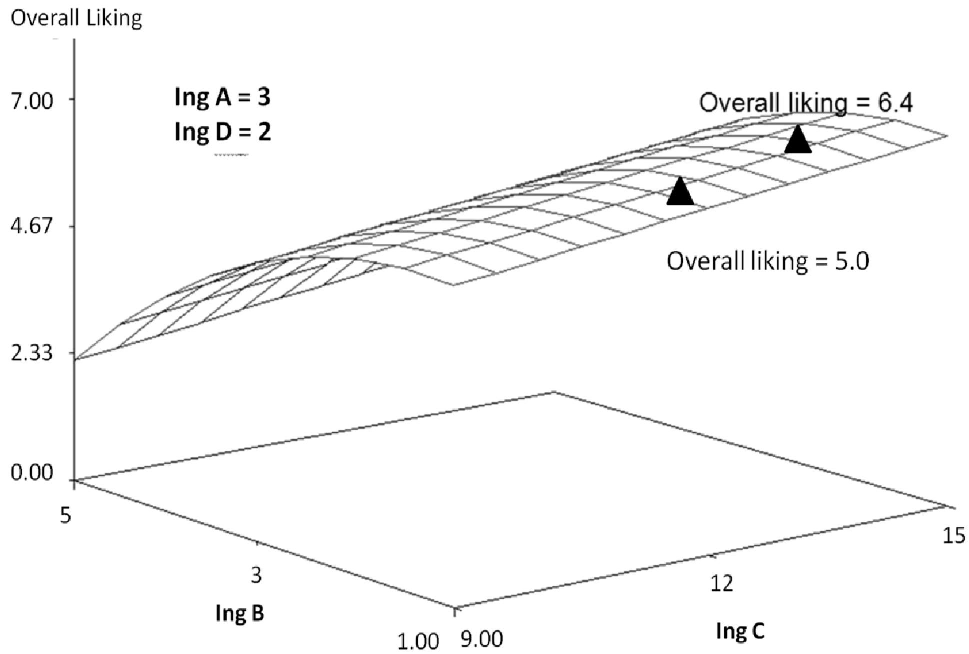
Overall Liking
Ing C = 14
Ing D = 4
7.00
4.67
2.33
0.00
4.00
2.50
Ing A
1.00
1.00
3.50
Ing B
6.00

(a)



Overall Liking
Ing B = 5
Ing D = 4
7.00
4.67
2.33
0.00
4.00
2.50
Ing A
1.00
9.00
12
Ing C
15

(b)



(c)

Figure 5: (a) SAS optimization contour map for A *vs.* B with Ing C and Ing D set constant and estimates of liking in the table portion of the output. (b) SAS optimization contour map for A *vs.* C with Ing B and Ing D set constant and estimates of liking in the table portion of the output. (c) SAS optimization contour map for B *vs.* C with Ing A and Ing D set constant and estimates of liking in the table portion of the output.

USER-DEFINED AND HISTORICAL DATA

It is costly to conduct a sensory study that involves descriptive analysis and consumer testing. From the business perspective, it makes sense that size and thus the cost of the experimental design should be modified to meet the budget requirement without sacrifice of research information. One can modify a standard design to meet research changes by simply deleting a row or rows in the design in consultation with a statistician.

A good strategy is to reduce the historical data to a simple response surface. Let's look at three such designs. Figs. **6–8** show us response surfaces with two variables. These are pentagon (Fig. **6**), the hexagon (Fig. **7**) and the octagon (Fig. **8**), respectively. They constitute three designs that can be easily implemented using the Historical Data option by the research and sensory analysts.

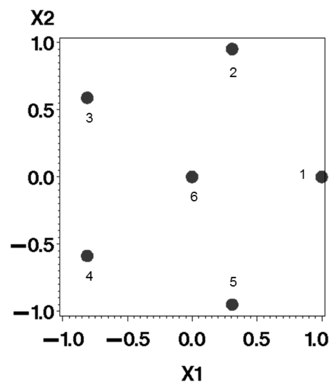


Figure 6: Pentagon design. Four levels of X1 (–0.809, 0, 0.309, 1) and five levels of X2 (–0.951, –0.588, 0, 0.588, 0.951).

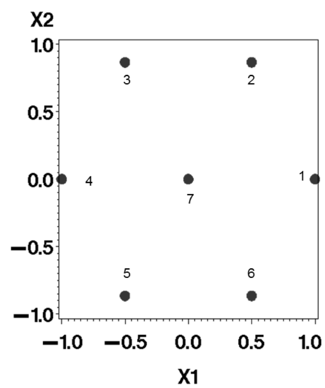


Figure 7: Hexagon design. Five levels of X1(–1, –0.5, 0, 0.5, 1) and three levels of X2 (–0.886, 0, 0.866).

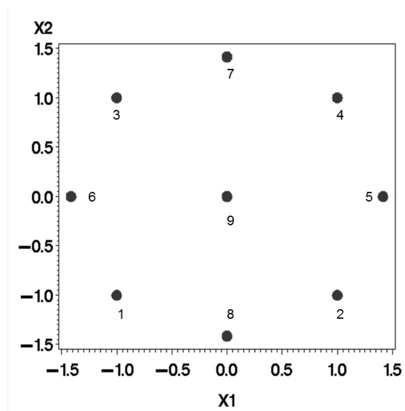


Figure 8: Octagon design. Five levels of each variable (–1.414, 1, 0, 1, 1.414).

In these figures, the units for X1 and X2 are coded or standardized levels. These designs belong to the well-known rotatable designs discussed in many textbooks (Cochran and Cox, 1992; Gacula and Singh, 1984; Gacula, Singh and Altan, 2009; Montgomery, 1991; Myers and Montgomery, 1995). They are important designs in optimization studies because of the role of the center point. In practice, optimizing an existing product by the addition or replacement of one or more ingredients uses the existing product as the center point for optimization. Thus, optimal areas can go in any direction from the center and it is expected that the predicted response has a similar variance.

The coordinates of the designs are coded factors that will be typed into the Historical Data option. They appear in Table **11** along with their respective response variables, *i.e.* sensory scores. In the actual experiment, it is recommended to replicate the center point more frequently than the peripheral design points. Such replication generates greater predictability of the response variables and thus a more reliable response surface map.

Table 11: Design Point Coordinates.

Design	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Pentagon									
X1	1.0	0.309	–0.809	–0.809	0.309	0			
X2	0	0.951	0.588	–0.588	–0.951	0			
Hexagon									
X1	1.0	0.500	–0.500	–1.0	–0.500	0.500	0		
X2	0	0.866	0.866	0	–0.886	–0.866	0		
Octagon									
X1	–1	1	–1	1	1.414	–1.414	0	0	0
X2	–1	–1	1	1	0	0	1.414	–1.414	0

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Cochran, W. G. & Cox, G. M. (1992). *Experimental designs* (2nd ed.). New York: Wiley and Sons, Inc.

Gacula Jr., M. C. (1993). *Design and analysis of sensory optimization*. Trumbull, CT: Food & Nutrition Press, Inc.

Gacula Jr., M. C. & Singh, J. (1984). *Statistical methods in food and consumer research*. Orlando, FL: Academic Press, Inc.

Gacula Jr., M. C., Singh, J., Bi, J., & Altan, S. (2009). *Statistical methods in food and consumer research* (2nd ed.). Orlando, FL: Academic Press, Inc.

Montgomery, D. C. (1991). *Design and analysis of experiments*. New York: Wiley and Sons, Inc.

Myers, R. H. & Montgomery, D. C. (1995). *Response surface methodology*. New York: Wiley and Sons, Inc.

SAS Institute Inc. (1999). *SAS/STAT user's guide*, version 8. Cary, NC: SAS Institute Inc.

Stat-Ease Inc. (1997). *Design-Expert® software, graphical version*. Minneapolis, MN: Stat-Ease Inc.

CHAPTER 10

Measuring Interest and Price for Sensory Experience: Application to Hotels

Johanna Fyrbjörk[1,*]. Rieko Shofu[2], Linda Ettinger Lieberman[3] and Howard R. Moskowitz[3]

[1]IPOS Sweden, Stockhom, Sweden; [2]Corporate Design Department, Hakuhodo Inc., Tokyo, Japan and [3]Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The RDE-based hotel study reported here shows that it is not the particular sense, but the experience that drives consumer interest and willingness to pay. A total of 315 respondents evaluated experimentally designed vignettes, comprising different combinations of positive, pleasant sensory experiences that a hotel might offer its guests as a point of differentiation. Each respondent evaluated a unique set of these vignettes. The ratings to the vignettes were deconstructed into the contribution of each sensory experience as a driver both of interest in the hotel and of relative amount of money one was willing to pay *versus* a standard one-night hotel cost. These experiences covered four different aspects of each of four of the five senses (seeing, touching, smelling and hearing). It is not the particular sense but the particular experience that drives interest and amount willing to pay. Three mind-set segments emerged: sensory seekers, fragrance and touch and design and relaxation.

Keywords: Sensory experience, price analysis, RDE.

INTRODUCTION

The five-sense experiences are deeply connected to people's consumption and purchase behaviors. We are only somewhat aware that these five senses unconsciously affect our decision-making at the time we make a purchase. However, the number of hugely popular products shows that appealing to the consumer's five senses through promotional activities has contributed to these products' success.

Take, for example, the coffee shop chain that is still expanding the number of stores even though several years have passed since the height of their popularity

***Address correspondence to Johanna Fyrbjörk:** IPOS Sweden, Stockhom, Sweden; Tel: 08 - 598 998 98 08-598998 98; Email: johanna.fyrbjork@gmail.com

Alex Gofman and Howard R. Moskowitz (Eds)

in Japan. These stores show how the five-sense experience may be applied in a commercial environment. When you enter the coffee shop and look around, you recognize that a lot of care has been taken with even the smallest details in the establishment, such as the displays and colors, background music, space brightness, chairs, the feel of the coffee cup, as well as the taste of the coffee. The consumer can have a unique five-sense experience that can only be provided at that particular coffee shop. These experiences can be subtle, such as putting holders on the coffee cups, to help consumers hold hot coffee. This care, conveyed through the senses, should certainly resonate with customers (Hakuhodo, 2006).

Vision, more than any other sense, has been a topic of a number of case studies. Take, for example, “colors”. Which brand actually strongly appeals to a customer’s color vision? The simplest method to identify that brand asks people whether they can recall a particular color for a specific brand name. People associate some brands with definite colors. It suggests that these brands have strong visual appeal/power that is strongly associated with specific colors in people’s minds and that brands and colors are inseparable companions. For example, one research study instructed people to name brands, as many as they could recall, with the requirement that the brand be attached to the color “red”. People recalled a variety of brand names such as an automobile maker, mobile-phone company, beverage maker, fast food restaurant and so on (Hakuhodo, 2006).

The same question should be posed to make clear whether the hearing sense also gives a strong impression of the brand. How about the start-up sound of a personal computer (PC), for instance? How about electronics retail stores that supply that type of PC? These brands should be associated with specific sounds as well as colors.

As for olfactory experience, smells and fragrances are seen as one of the many effective tools for branding. The automobile industry has conducted research and development regarding smells as well as sounds. Some automakers make attempts to connect smells to their cars by adding perfume, blended with the adhesive used to apply the car’s interior material, which is then identifiable as the car brand’s original perfume. These attempts have built ties between customers and the brand through the memory of smell. Some car enthusiasts can even identify the automaker by the scent of the car!

Sensory experience can be used to drive brand personality. Given the increasing importance of sensory experience to product success or failure, a question that never seems to be addressed head-on is how to measure the true value of the experience and specifically, how to value sensory experience. Is a pleasant auditory experience judged to be far more or less expensive than a pleasant olfactory experience, or kinesthetic/touch experience?

Trying to place a value on the sensory experience may seem, at first, a bit remote, a bit theoretical and academic. Yet, despite the intangible nature of sensory experience, we do know that people pay money to hear opera, to smell fragrance, to see art and so forth. So, there is a value attached to the experience. The empirical question is how to put a numerical value on these sensory experiences. In order to attach economic considerations to the measures, we measure both interest in the experience (an attitudinal measure) and amount of money a consumer is willing to pay (a surrogate econometric measure).

Experimental psychologists, specifically psychophysicists, measure sensory experience and relate it to physical measures. Once the bonds of psychophysics are relaxed, one can work with *descriptions* of these experiences, as marketers do, measuring the interest in such experiences. This advance over traditional sensory science can be made even more powerful by attaching economic considerations to the measures. The investigator can now measure both interest in the experience (an attitudinal measure) and amount of money willing to pay (a surrogate econometric measure).

The study of sensory experience becomes even more interesting and productive when, in turn, the test stimuli are systematically varied according to an experimental design (through conjoint analysis), so that the stimuli comprise several stimuli conjoined in a vignette. Consumers rarely experience one message. Marketers do not promote only one feature of a product or service, such as a hotel, but rather present a combination of features in an advertisement. In such combinations, the different elements compete with each other to drive the response. When these elements compete, it is impossible for the respondent evaluating the combination to be "politically correct". In a study of several vignettes, where these vignettes change rapidly, one after another, the respondent relaxes and evaluates the vignette as a whole, much as one evaluates advertising and real-world offerings. In such cases, the truly best elements emerge after the deconstruction of the response into the contribution generated by each element in the vignette.

What to Look for in Designed Experiments

We approach the issue of sensory economics in a systematic way. Our objective is to put numbers onto the desirability of sensory experiences, such that these numbers reflect the private wishes of the respondents. By presenting the experiences as a vignette, we force the consumer respondent to integrate the sensory information presented in the vignette. We then use standard statistical procedures, typically ordinary least-squares regression analysis (OLS), in order to estimate the part-worth contribution of every element. When each respondent evaluates a variety of different vignettes and when the elements in the vignettes vary independently of each other, OLS estimates the contribution of each element, even when the respondent cannot articulate at a conscious level what is important. The same approach applies to dollar value. When the respondent rates the dollar value of the vignette, or some equivalent (*e.g.,* percentage of the standard price they are willing to pay), OLS estimates the dollar contribution of each element. Often respondents faced with separately estimating the dollar value of each part of a vignette find the task to be onerous. Yet when the elements are combined into a vignette and the respondent need only rate the combination in terms of dollars, the task becomes remarkably easy to do. Indeed, consumers do this every day when shopping for goods and services, so we should not be particularly surprised when they do it in a rule developing experimentation (RDE) study.

METHOD

Consumer respondents are invited to participate by an e-mail invitation. They read an orientation screen (Fig. **1**).

We worked with 20 elements, allocated to five silos, each with four elements. The five silos are the hotel name, then four elements each for touch, sight, smell and sound, respectively.

The elements are mixed and then evaluated in small, easy-to-read combinations of the elements (Fig. **2**; Box, Hunter and Hunter, 1978). Every respondent evaluates each element three times, in totally unique sets of combinations. By having every respondent evaluate totally unique combinations of the same element, it becomes possible to avoid any possible bias due to unsuspected interactions among elements that could bias the results. It further becomes possible to partial out all pairwise interactions among elements of different silos, an analysis that we will do at the end of this paper, when we look at the interaction between hotels and sensory offerings.

Dear Hotel Guests,

We are conducting a survey on **Hotels** to include/improve features for your comfort and convenience.

You can help us create the best hotel. When you read the full hotel description on the screen ... think about the following:

How interested are you in staying in this hotel?

1 = Not at all interested ... 9 = Very interested

Versus its typical, CURRENT 1 night **WEEKEND** stay ... how much should the hotel charge for this room?

1= 40% less, 2= 30% less, 3= 20% less, 4= 10% less, 5=the same, 6=10% more, 7 = 20% more, 8 = 30% more, 9=40% more

Please take your time and go through the study. Your insights are important to us.

Thank you!

Figure 1: The orientation page.

Themed hotel restaurant coordinated with the decoration of the lobby

Soundproof room offers quiet stay

Bathroom ambiance included with an aromatic effect for relaxation

Showers have an aromatic steam option to stimulate relaxation…the ultimate experience

Figure 2: Example of a test concept.

The elements appear independently of each other (Moskowitz and Gofman, 2005). In some cases, the concept is absent one or two sensory silos. This "incompleteness", while seeming a problem, actually is not a burden to the respondent who simply does the evaluation. The incompleteness permits OLS to estimate the absolute contribution of every element. The property of absolute value will become important when the RDE results are used in connection with creating a database of the consumer mind. In turn, this database will become increasingly valuable when the results from one study are combined with the results from other studies, with other elements, done in different places and at different times.

RESULTS

We begin by estimating the part-worth contribution of each element. The analysis is fairly straightforward, but it involves one simple transformation. We focus on the whether or not a respondent is interested in the sensory experience described by the concept. Although the respondents used a 9-point category scale, the convention of consumer researchers is to look at more absolute judgments: interested or not interested. Following this "all-or-none" way of looking at the data, we recode ratings 1–6 to "0" in order to represent "not interested". We recode ratings of 7–9 to "100" in order to represent "interested". Then we run the OLS, relating the presence/absence of the elements to the binary (0/100). The result is an "interest model". For this analysis, we combine all the data into one large file and run one general regression. We do not show the contributions of the four hotel names to interest, although they were included in the actual regression modeling.

The same deconstruction analysis can be done for the ratings of the amount one would pay. The analysis is straightforward. We begin by replacing the numerical rating of price with the percent that a person is willing to pay. Thus paying 20% more is replaced by the value 120, *etc.* The analysis then proceeds by OLS, estimating the part-worth contribution of each of the 16.sensory phrases, as well as the four hotel names (latter not shown). These results appear in the first data column in Table **1**. We discuss the results in detail in the following paragraphs.

What Interests a Prospective Customer?

The additive constant measures basic interest in the hotel. Of course, all respondents evaluated concepts comprising elements and were never asked to state just how interested they would be in a "hotel" without anything. Nor, in fact, do we believe that people can really articulate that notion of "basic interest". Yet, using statistical methods (especially regression) one can get a sense of this basic interest, or at least some aspect of it, by virtue of the additive constant. When we accept that additive constant as a measure of basic interest, Table **1** tells us that without any further specification of the hotel experience, there is only modest basic interest in the hotel. The additive constant is 24, meaning about 24%, or one person in four, would be interested.

Table 1: How Elements Drive Interest (% Rating 7–9 on a 9-Point Scale) and Relative Value (% Willing to Pay for the Sensory Experience, Above the Standard Room Rate).

Sense Element	Total	Interest Impact	Price %
	Constant	**24**	**93**
Feeling 1	Rich, lush and soft.uniquely crafted bed and pillows in your room for a comfortable sleep	14	3
Feeling 3	Rooms equipped with a massage chair	13	4
Hearing 4	Soundproof room offers quiet stay	12	3
Feeling 4	Showers have an aromatic steam option to stimulate relaxation.the ultimate experience	11	3
Smelling 2	All rooms equipped with an air purifier that has an aromatic function	9	2
Smelling 3	Bathroom ambiance included with an aromatic effect for relaxation	7	2
Hearing 3	Sound on demand system in every room.choose from a variety of music	5	2
Feeling 2	Natural selected linens and towels with pleasant texture and colors	5	2
Seeing 2	The lobby elevator color schemes are carefully selected for peace and comfort	4	1
Seeing 3	Hotel room color schemes carefully selected for peace and comfort	4	1
Hearing 1	Unique background music in the lobby developed.for a warm welcome	4	1
Seeing 1	Know what to expect inside just by looking at the exterior of the hotel	3	1
Seeing 4	Themed hotel restaurant coordinated with the decoration of the lobby	2	1
Hearing 2	Unique background music in each room.for a comfortable sleep and fresh awakening	2	1
Smelling 1	At lobby, guests welcomed by a pleasant fragrance	1	0
Smelling 4	Hotel features a “Fragrance Bar”.test many fragrances.choose best fragrance for your room	–1	0

The real story comes from the elements. The elements are the specifics of the hotel. We tried to write the elements in such a way that each element painted a word picture. From the RDE study and from the regression analysis, each element generates an associated utility. The utility, in turn, is the incremental percent of respondents who would rate the concept as 7–9 (*i.e.* be interested in the sensory experience) if the sensory experience were to be part of the concept. When the utility is negative, as it is for a number of elements, the effect is negative, subtractive. Putting the element into a concept about hotels actually decreases the percent of respondents who would rate the concept 7–9 on the 9-point scale.

Specific sensory experiences drive interest, but not all of them. Paint a word picture of the kinesthetics and touch and you're likely to get people interested. The notion of "rich, lush and soft…bed and pillows" is very strong, with a coefficient of +14. We interpret that +14 to mean that an additional 14% of the respondents would be interested in staying at the hotel (rate the concept 7–9) if the hotel were to feature the richly crafted, lush and soft bed and pillows. The same goes for a massage chair, soundproof room and showers with an aromatic steam option.

Not everything works, however. Feature a fragrance bar and no one is interested, at least based on the average.

Price Premium

Let's move beyond interest to the price a person is willing to pay and in turn, the contributions of the different elements. We follow the same logic, *i.e.* use regression. There is one major difference, however. When it comes to price, our respondents had to select a rating that in turn corresponded to a price. We have to move away from the actual rating and move to the actual price as a dependent variable. We can do that quite easily by replacing the price statement by the relative amount. We will deal in percents. Paying the current price is thus 100, corresponding to paying 100% of current. Paying 20% less is tantamount to paying 80% of the current price and therefore the value is 80 as a dependent variable.

With this approach in mind, let's look at the results of our regression analysis. The independent variables were the elements; the dependent variable is the percentage a consumer would pay *versus* today's price (pegged at 100%).

1. Looking at the right-hand column of Table **1**, we should be struck by the fact that *people do not want to pay for what they get*. The additive

constant is 93, meaning that without any elements, on the average people are willing to pay about 93% of the standard room rate.

2. The elements that are most interesting are also those that the respondents would pay extra for, but the key here is only *slightly extra*. We're talking about 3–4% extra over the normal room rate.

3. *Homo economicus*, economic man or economic considerations, are far more conservative than interest. One might make the wrong decision by looking at interest alone. Interest values can swing far more positive and lead to false expectations. Tacking on amount of payment as a second rating question shows the strength of the element in far greater clarity.

Different Strokes for Different Folks—The Role of Mind-Set

Thus far, we have dealt with the data as if the respondents comprised one homogeneous group of prospective customers. The reality of the matter is that people differ from each other, sometimes in small ways, sometimes in large ways. We are not talking here about the demographic differences in gender, market, or income. Nor are we talking about the so-called behavioral differences that we can measure today with tracking systems that show differences in Web behavior. Rather, we are talking about more profound, deeper and structural differences in what people truly like. Marketers recognize that people differ profoundly. Over the years and responding to these profound differences, marketers and researchers created a variety of psychographic testing systems by which they could assign people to neat and tidy buckets. People in a bucket were presumed to share the same values. The goal was to market to these individuals in a similar manner, because, the thinking was, individuals in the same psychographic group should have the same mind-set to many different categories of offerings.

The general psychographic approach did not work because at a granular level, at the level of concrete, specific experience, people differ profoundly, even when the people fall into the same general psychographic group. There's just too much granularity and specificity in the world to allow such general segmentation methods to work for the particular issues of the types that marketers face in their daily effort to convince consumers to buy. In other books and papers (Moskowitz, Poretta and Silcher, 2005), one author (HRM) has proposed a much simpler approach. The approach posits that there are individual mind-set segments in any particular product and service area, including, of course, hotel preferences. An effective strategy to cope with these granular-level segments is to discover them at the time the marketing/sales opportunity arises, rather than hope that a general segmentation would translate to the specific opportunity. That is, there is not necessarily a general segmentation needed, but rather a strategy to understand local, momentary segmentation and then take advantage of that segmentation for the particular opportunity at hand.

There is a bit more to this notion of granular-level segmentation. That "more" includes the methods for understanding the segments and the practical application of methods to use the segmentation for business purposes.

At a practical level, the segments for a particular product or service are distributed in the population but are hard to discover with standard data-mining methods. They may or may not be discoverable by knowing a lot about the respondent. A better way is to use a simplified active intervention test, something specific to the topic being investigated. (A good analogy is sensitivity testing by allergists for allergies, which is done by a specific-level intervention test, the scratch test).

At the operational level, these segments can best be revealed and their nature understood through a short "test". That approach presents the individuals with test stimuli, such as experimentally designed vignettes (concepts), obtains their responses (our 9-point ratings) and then builds a model for each individual, showing how the individual elements "drive" the response. Then, individuals are clustered together, based upon the pattern of utilities or coefficients in this model. The foregoing approach to discover the segments for a specific product or service does not require that individuals who fall into the same cluster or segment be linked together for other products or services.

Following this notion of dividing people by the pattern of their responses at the level of specificity for hotels, we segmented the 315 respondents, based upon the pattern of coefficients from the regression equation (see Table **2**). The segmentation is a statistical operation. We should note that the job of the investigator is to select the number of segments and to assign a name to each segment. The actual segmentation method, called k-mean clustering, is well-defined, objective, part of most statistical packages dealing with data analysis and outside the control of the investigator. The operating rule is that there should be as few clusters as possible (parsimony) and that the elements in the cluster should tell a simple story that convinces (coherence).

Table 2: Percent of Respondents Interested in the Hotel, Based on the Elements in the Vignettes and Relative Amount Each Segment (S1–S3) Would Pay for the Most Important Sensory Experiences for That Segment. Bold Numbers Denominate Positive Impact on Interest.

	Interest in Hotel Based on Element			**Relative Price Willing to pay (% of Current Price that One Would Pay)**		
	S1	S2	S3	S1	S2	S3
	84	147	84	84	147	84
Additive constant (basic response, without elements)	29	20	26	94	92	95
Segment 1—Sensory Luxury						
Rooms equipped with a massage chair	**15**	11	13	**4**	5	2
Showers have an aromatic steam option to stimulate relaxation.the ultimate experience	**13**	11	10	**4**	4	2
Rich, lush and soft.uniquely crafted bed and pillows in your room for a comfortable sleep	**13**	14	14	**3**	4	2
Soundproof room offers quiet stay	**10**	16	8	**3**	4	2
Segment 2—Fragrance and Touch						
Soundproof room offers quiet stay	10	**16**	8	3	**4**	2
All rooms equipped with an air purifier that has an aromatic function	–4	**15**	10	0	**3**	2
Rich, lush and soft.uniquely crafted bed and pillows in your room for a comfortable sleep	13	**14**	14	3	**4**	2
Bathroom ambiance included with an aromatic effect for relaxation	–2	**14**	5	–1	**3**	2
Rooms equipped with a massage chair	15	**11**	13	4	**5**	2
Showers have an aromatic steam option to stimulate relaxation.the ultimate experience	13	**11**	10	4	**4**	2
Hotel features a "Fragrance Bar" test many fragrances.choose best fragrance for your room	–11	**11**	–11	–2	**3**	–1
At lobby, guests welcomed by a pleasant fragrance	–7	**9**	–4	–2	**2**	0
Sound on demand system in every room.choose from a variety of music	7	**8**	–1	2	**3**	1

Table 2: cont….

Segment 3—Design and Relaxation						
Rich, lush and soft.uniquely crafted bed and pillows in your room for a comfortable sleep	13	14	**14**	3	4	**2**
Rooms equipped with a massage chair	15	11	**13**	4	5	**2**
The lobby elevator color schemes are carefully selected for peace and comfort	2	1	**11**	1	0	**1**
Hotel room color schemes carefully selected for peace and comfort	1	2	**11**	1	1	**2**
Showers have an aromatic steam option to stimulate relaxation.the ultimate experience	13	11	**10**	4	4	**2**
All rooms equipped with an air purifier that has an aromatic function	–4	15	**10**	0	3	**2**
Themed hotel restaurant coordinated with the decoration of the lobby	1	–2	**9**	1	1	**1**
Know what to expect inside just by looking at the exterior of the hotel	2	1	**9**	0	1	**1**

Our segments do not fall into nice, neat, or tractable patterns. Nor do people. The results suggest three segments, two small and one large. Let's first look at the interest values, *i.e.* the element utilities on the left side of the table. These numbers tell us the basic interest in staying at the hotel and the contribution to the basic interest made by each sensory feature that the hotel can offer.

1. The first segment, comprising 84 of the 315 respondents, is modestly interested in the hotel stay (additive constant of 29, meaning that without anything else, only 29% of the individuals in this segment would rate a hotel concept 7–9). What's important here is luxurious relaxation. Relaxation comes in the form of the rich, lush and soft bed and pillows; the message chair, even the soundproof room and the aromatic steam option. But do not try to attract this first segment with fragrance; it does not work, except when the fragrance is embedded in the shower and only to promote relaxation.

2. The second segment, by far the biggest, comprises 147 of the 315 respondents. They are not really interested in the hotel, with an additive constant of 20, the lowest of the four segments. To get these individuals the hotel must offer specific sensory experiences. This

second segment wants "sensory experience," primarily fragrance and touch. They really like not only the lush pillows, the massage chair, but also the soundproof room. They want touch; they want smell; and they want a quiet room.

3. The third segment, comprising 84 respondents, responds to design and relaxation. Unlike the other segments, they are visually oriented.

4. What we see from the segmentation is that the hotels should offer several experiences. These experiences must be specific, not general. The research tells the hotel what experiences will attract and who will be attracted. The mind-set segments are not opposite to each other, but rather comprise people who react to the positives, but to different degrees.

5. Not everything is good. The hotel can make a mistake. It is possible to turn off these segments, not only primarily with fragrance, but also with sound. Thus fragrance could turn out to be a polarizing factor.

But Will They Pay?

In Table **1**, showing results for the total panel, we saw that *homo economicus* was very conservative. The large effects that we observed for ratings of interest disappeared, to be replaced by rather small changes in the relative amount of money one would pay. That makes sense, since interest is not the same thing as putting out money.

But what about the segments? Our segments show strong responsivity to the appropriate sensory features that a hotel would offer. Yet, will our segments pay for what they like? Or do we see the same conservatism within our segment results? The simple answer is "yes", the conservatism remains. We do not get a sense of wide swings in the price consumers are willing to pay. In contrast, there are a number of sensory features that might each command an additional 4–5% in the price one is willing to pay.

Fig. **3** suggests that prospective guests are likely to pay somewhat more for sensory experiences that they like, no matter the segment to which they belong.

Thus it pays to give the prospect a positive, desired sensory experience; it's more likely that they'll pay more, just not as much more as one might hope!

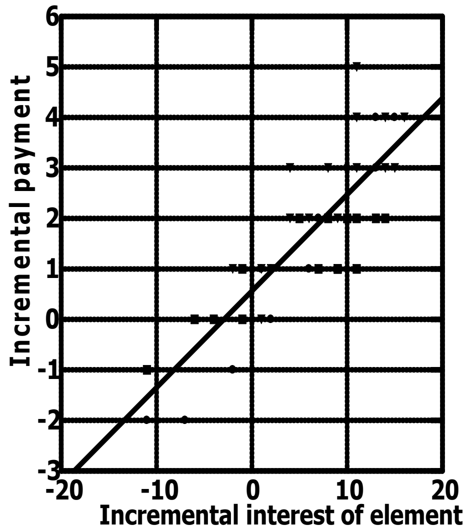


Figure 3: Relation between element utility (abscissa) and incremental/decremental payment (in %). Each point on the plot corresponds to an element and a segment. The plotted data show the interest–payment relation for all three segments and for the 16 sensory experiences, or a total of 48 points.

Do Hotel Chains "Interact" With Desired Sensory Experiences?

Up to now we have been dealing with sensory experiences independent of hotels. Although we worked with 20 test elements in our vignettes, we have concentrated only on the 16 sensory elements and not paid attention to the four hotel names. These four appear in the model, totally independently of the 16 sensory elements.

The reality of today's world is that hotels attempt to create an "image" through advertising in mass media and through the experiences they create when the guest actually stays in one of their rooms. One of the most important questions is whether a specific sensory experience "goes with" a hotel.

What do we find when we search for synergistic combinations of a hotel name and sensory experience? Here is how to do the search and find synergies and suppressions that might be lurking in the data, totally unbeknownst to us. The strategy is particularly important when we deal with brands and sensory experience.

1. We begin by stratifying the 7,875 different combinations that respondents evaluated. As there are 315 respondents, each of whom evaluated a relatively unique 25 combinations, the total number of combination is 7,875.

2. Let us sort those into the concepts having no hotel name and the four layers of concepts, with each of the actual hotel names (Ritz-Carlton, Hilton, Marriott and Sheraton).

3. For each of these five "layers" or strata, we run the same model relating the presence/absence of the 16 sensory experiences *versus* interest in staying at the hotel.

4. Putting the analysis in perspective, we want to discover whether adding the name of the hotel to the basic sensory experience changes the desirability of the experience.

5. The analysis is straightforward by OLS. We merely run the model with each of the five separate strata.

6. Our results appear as the five data columns in Table **3**. The general pattern seems to be *slight suppression of sensory experiences by the different brand names*. In the absence of a hotel brand name, the strongest sensory experiences do quite well, as we see in the first data column labeled "No Hotel Name". For the three strongest elements (massage chair; soundproof room; rich, lush and soft bed and pillows), combining them with hotel brand name somewhat diminishes their impact. This drop is even more dramatic for the less powerful sensory experiences, such as an air purifier with an aromatic function. The hotel name diminishes the impact of the sensory offerings.

7. The bottom line here is that putting a hotel brand name on the offering reduces the impact of the sensory experience. However, we do not know whether this reduction comes from a clash between the hotel brand name and the sensory experience, or whether brand names simply suppress the contribution of the elements.

Table 3: Performance of Strongest Sensory Experiences in Concepts With No Brand Name and in Concepts With One of Four Well-Known Hotel Brand Names.

	No Hotel Name	**Ritz Carlton**	**Hilton**	**Marriott**	**Sheraton**
Constant	19	22	26	26	28
Rooms equipped with a massage chair	16	10	13	11	12
Soundproof room offers quiet stay	14	15	12	9	8
Rich, lush and soft.uniquely crafted bed and pillows in your room for a comfortable sleep	13	15	18	9	15
Bathroom ambiance included with an aromatic effect for relaxation	12	7	4	9	3
Showers have an aromatic steam option to stimulate relaxation.the ultimate experience	12	10	14	9	12
All rooms equipped with an air purifier that has an aromatic function	10	12	5	8	6
Hotel room color schemes carefully selected for peace and comfort	9	4	1	4	3
Sound on demand system in every room.choose from a variety of music	7	6	5	4	5

DISCUSSION AND CONCLUSIONS

Research using text questionnaires gets to the heart of the mind; what is important. Without the tools of experimental design and conjoint analysis, one could not really know what is important. One might show the prospective customer pictures of rooms and get responses. Yet, it is apparent that such pictures are by their very nature incomplete. There are the other senses to consider. These senses cannot be presented to the prospective guest other than by words. And, furthermore, we see that the visual aspect of the hotel is only one part and a relatively small part, in terms of what drives consumer interest.

The important next step is to design the communication piece. How does the hotel communicate fragrance, lushness of the pillows, special showers with aroma, massage chairs? It may be possible to communicate some of these benefits and attractants by pictures. Otherwise, these features will have to be wordsmithed for advertisements. The communication of such benefits is not part of this study but is a natural next step.

Beyond communication is the use of the data to create the offerings. By itself, the study here provides guidance only in text form. There are no pictures and even if there were, pictures work primarily for visual design. They may communicate equipment, but they cannot take the place of sensory experience. It is at sensory experience where we stop in this paper, waiting for the next installment, to be contributed by psychophysics.

We began this chapter with a systematic exploration of how to understand what sensory experiences are valued by the hotel customer, both in terms of interest and then in terms of willingness to pay. Our analysis stopped at identifying what types of experiences, on a described basis, seem to work best, how people differ from each other in mind-set segments. This approach is conceptual, dealing with the psychophysics of the mind, of evocative description. We have to move forward now.

Experimental design, not only of concepts, but also of actual experiences, of actual stimuli arranged in different combinations, will provide new opportunities for business and for design engineering. And it is there that we will encounter the new frontier, where visions and opportunities will emerge and where they will be realized as the science of the mind takes one of its next steps.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Hakuhodo Brand Design. (2006). How to create brand personality—practice of five senses branding. Tokyo: Diamond Inc.

Box, G. E. P., Hunter, J., & Hunter, S. (1978). Statistics for experimenters. New York: Wiley and Sons, Inc.

Moskowitz, H. R. & Gofman, A. (2005). System and method for performing conjoint analysis. U.S. Patent Pending, US2005/0177398, 2004.

Moskowitz, H. R., Porretta, S. & Silcher, M. (2005). Concept Research in Food Product Design and Development. Ames, IA: Blackwell Publishers.

Stevens, S.S. and Stevens, G. (1975). Psychophysics: An introduction to its perceptual, neural and social prospects. New York, NY: Wiley and Sons, Inc.

PART III

MESSAGE OPTIMIZATION

CHAPTER 11

Messaging Across National Markets—Effectiveness and Segmentation

Gillie Gabay[1], Samuel Rabino[2] and Howard R. Moskowitz[3,*]

[1]College of Management Academic Studies, Rishon Letzion, Tel Aviv, Israel; [2]Northeastern University, Boston, Massachusetts, USA and [3]Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The object of the study is to identify effective communication messages for consumers facing product/brand options, when the category is a packaged good, in a competitive and saturated category in different countries. The challenges facing multinationals include differentiating their brands from competitors and creating brand equity. Segmentation determines groups of individual consumers similar in preferences and in intended buying behaviors. These segments may transcend national boundaries. Identifying such segments simplifies the design of advertising messages that appeal to audiences regardless of their country of origin and reduces promotional costs. This chapter presents ways to identify effective communication messages for consumers in the toothpaste category for three developing countries with developed markets. The study determines whether traditional product segmentation strategy based on attributes (forms, flavors and ingredients) could be augmented by a mind-set segmentation. The results suggest three key segments that transcend the countries and provide a basis for a successful communication strategy. The study's major contribution is the delineation of a framework for data assembly and identification of metrics measuring different aspects of consumption patterns in a highly competitive packaged goods category.

Keywords: Consumer segmentation, rule developing experimentation, conjoint analysis, message optimization.

INTRODUCTION

Many fast-moving consumer goods are exhibiting slow growth in developed and emerging economies as well as in saturated markets. While these products may become commodities, competitors may find new product categories that have not been exploited. Thus, it is impossible to abandon these products to the competition. Such opportunities to find new categories may exist and may provide

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: (077) 700-1591; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

a competitive advantage, but each day they recede as the world of fast-moving consumer goods becomes smaller. Globalization of brands, hyper-competition in market after market and the need to maintain corporate profitability all drive the need to identify better messaging as one way to maintain market share. Thus, the ability of multinational corporations (MNCs) to design effective promotional messages for today's competitive landscape has become an important source of competitive advantage that entails smaller risks than those entailed by the development of new categories.

In this exploratory study, we present a knowledge-building approach using rule developing experimentation (RDE) that marketers use to identify selling propositions and in so doing discover segments that transcend national boundaries. The theoretical foundation of this study is the abundant research conducted in the areas of international marketing orientation—standardization *versus* differentiation related issues—and the study of alternative international segmentation schemes. We introduce the empirical approach by first discussing those aspects of these two fields of research that are particularly relevant for this study.

The issue of marketing uniformity and especially product standardization *versus* adaptation to the idiosyncratic circumstances, has been analyzed in depth (*e.g.,* Szymanski, Bharadwaj and Varadarajan, 1993). Proponents of standardization argue that the internationalization of companies, the growth of international communication channels and media, the Internet, travel activities of consumers, *etc.* all produce a converging world to which the marketer must adapt (Theodosiou and Leonidou, 2003; Papavassiliou and Stathakopoulos, 1997). Customers in distant parts of the world increasingly tend to exhibit similar preferences and demand the same products (Jain, 1989; Levitt, 1993). Not everyone agrees, however. Other research has cautioned against an unconditional acceptance of this thesis (Boddewyn, Soehl and Picard, 1986; Douglas and Wind, 1987; Whitelock, 1989). More recently, Bain and Company consultants argued that the era of standardization is ending—consumer communities are growing more diverse in ethnicity, wealth, lifestyle and values (Rigby and Vishwanath, 2006). Either way, from the perspective of resource allocation and the need to properly communicate the quality points of the core products, consumer packaged goods marketers must craft their messages to consumers very carefully (Barwise and Farley, 2005).

Segmentation studies in the international context are based on the premise that companies should attempt to identify consumers in different countries who share similar needs and desires. Marketers are often faced with the dilemma of deciding whether to segment markets on a country-by-country basis or instead opt for a global marketing strategy, which targets one or more similar segments that exist in several countries (Broderick, Greenley and Dentiste Mueller, 2007).

The key action to be taken is message creation. Embedded in the dichotomy of standardization *versus* differentiation is the need to create effective advertising messages. Effective messaging is even more critical for those situations where consumer behavior and attitudes transcend national borders (Ter Hofstede, Steenkamp and Wedel, 1999). Message decisions (also referred to as creative strategy) focus on those specific unique selling propositions (USPs) that should be communicated. The decisions must take into account the specifics about what the communication is intended to achieve in terms of consumer behavior. These decisions have important implications for the choice of advertising medium or execution as certain media can better accommodate specific creative requirements (use of color, written description, product benefits, *etc.*).

This chapter presents an empirical knowledge-building approach to identify which communication or message components are appropriate to specific mind-set segments. Segments in this context are defined by their responsiveness to an alternative set of communication messages. It should also be noted that messaging research, which is a subset of the current body of knowledge dealing with international advertising issues (*e.g.,* Fastoso and Whitelock, 2007), was characterized as providing little insight into the use of effective advertising message characteristics (Stafford, 2005). This study counters that complaint and provides a new way for marketers to use messaging research as a key contributor to both tactical actions and strategic information, respectively.

UNDERSTANDING CONSUMER DYNAMICS IN DEVELOPING COUNTRIES

The three countries represented in this study are also members of the Asia–Pacific Economic Cooperation (APEC). The economies of Australia, an industrial market and the Philippines, a developing economy, exist in a relatively close geographic proximity and they have a common spoken language (English). Mexico is another developing economy, located in an entirely different region. Mexico is exposed to the same group of multinationals competing for a share in the same category. Considering the exploratory nature of this study and having access to data collection outlets in Mexico, we included Mexico in this project.

Toothpaste was chosen as the focal product class due to its high purchase frequency and usage. Toothpaste is also a good product to study because it is familiar and brands are recognized in the mega world of fast-moving consumer packaged goods. The brand evaluated in this study is sold in over 100 countries worldwide.

Over the past three decades, toothpaste evolved from a few simple varieties to a large number of entries. The international acceptance of toothpaste has clearly grown, but unlike the knowledge-base of consumer choice in the USA or Europe, not much is known about the dynamics of consumer preference in developing countries. Couple this paucity of knowledge with the need of businesses to create and market the next generation of toothpaste and an opportunity to study the dynamics of responses in developing countries emerges. In these developing countries, the dynamics of consumer response to new, easily affordable products are often driven by a combination of messaging, consumer experience and increasing awareness of health. In developing markets, the growth of the category has been slowed due to intense competition, erosion of margins and crowded brand extensions (Information Resource Inc., 2007). In developing markets, purchases of toothpaste brands are often hampered by their price relative to per capita income as well as high levels of inflation (Czinkota and Ronkainen, 2007). Furthermore, oral hygiene awareness and per capita consumption of sugar and tobacco makes advertising messaging in these markets particularly challenging (De Paola, 2007).

STANDARDIZATION *VERSUS* COUNTRY-BY-COUNTRY INDIVIDUALIZATION

Previous research evaluated the relation between standardization and advertising, focusing on antecedent factors (*e.g.,* cultural environment, market diversity) as key variables that drive the choice between a standardized strategy and an adapted one (Papavassilliou and Stathakopoulos, 1997). More recently, Wei and Jiang

(2005) suggested that whether international advertising practices are standardized or localized depends largely upon differentiating a creative strategy (selling propositions) from its execution (media related decisions). The Wei and Jiang study reported that advertising campaigns are more likely to be standardized at first when the creative strategy is formed, but then localized when executing that strategy in the different countries/markets. In general, studies of the differentiation between creative strategy and execution suggest that creative strategy (the message containing themes or selling points) can be standardized, whereas execution formats are usually adapted to the unique environment of different local markets (Duncan and Ramaprasad, 1995).

While ample evidence substantiates the practice of many multinationals in designing different message appeals for use in different countries, the topic of just *how* consumers in different countries respond to different appeals (or combinations of appeals) has received less attention (Aaker and Williams, 1998). As a result, advertisers are not always clear as to what extent advertising messages or themes can be internationally standardized. Some research has been conducted in the area of message framing and the determinant of message framing effects (Orth and Firbasova, 2003).

The topic of investigating possible segmentation schemes while addressing the likelihood of behavioral homogeneity and heterogeneity has been examined elsewhere (*e.g.,* Broderick *et al.* 2007). In this study and chapter, we offer a strategic perspective of alternative international advertising messages based on the evaluation of the same messaging within countries and across national boundaries. With the increased competition in the global marketplace and increased international immigration, there is a critical need to identify viable international segments and reach those segments with effective messaging that aids in positioning and selling products across national borders. Such capabilities will help brands survive. In other words, articulating the relevant *core benefits propositions* of a brand for the appropriate segment is more likely to increase responsiveness to the message.

INTEGRATING IT ALL—BRAND EQUITY, SEGMENTATION AND POSITIONING AND BUILDING INTERNATIONAL BRANDS

In the case of multinationals that are active in the promotion and selling of consumer packaged goods, the need to establish and sustain brands in various international markets is critical. To that end, the approach that identifies potential customers along with the most effective communication will become the foundation of a potential marketing strategy.

When tourism, migration and information technology each contribute to familiarizing populations with the plethora of brands, appropriate messaging will become ever more powerful. The messaging will lower the mental effort and search time needed by the consumer to invest in choosing a product and in turn will reduce the risk of the decision, at least as the consumer perceives that risk (Avery and Keinan, 2008).

The effect of a good brand strategy, supported by messaging, will be the creation of brand equity. In turn, brand equity will drive a revenue premium as well as create cost savings. Multinationals with strong brands gain efficiency in marketing costs because consumers are more *receptive* and responsive to the brand messages. The same message could aid as multinationals attempt to expand beyond their current base in one country market to other international markets by extending a brand into new markets or categories (Avery and Keinan, 2008).

In their quest for competitive advantage, marketers create brand knowledge structures, *i.e.* the understanding of product offering and/or information about these products (Keller, 2001). When deployed in a multination market, the communication strategy should associate a brand with other people, places, concerns, or even other brands. This strategy builds knowledge that might otherwise be difficult to achieve through other, more direct product marketing programs. Building and understanding a consumer brand-knowledge structure, including the ability to understand the differences among brands, becomes crucial for effective communication (Keller, 2001).

THE SCOPE OF THE RDE EXPERIMENT—MESSAGES DRIVING RESPONSES

The aim of this exploratory study was to learn more about which messages and product benefits/features drive acceptance in two developing markets (the Philippines, Mexico) and one developed market (Australia). These three markets represented different usage and consumption patterns. They were selected as test areas to evaluate alternative messaging combinations. A similar three-country study examined performance drivers in the globally focused marketing organizations (Hult *et al.* 2007). In the same vein, a recent cross-national study evaluated lifestyle segments for the fashion industry in Korea, Europe and the USA (Ko *et al.* 2007). What about *our* cheese messaging study in Europe?

The specifics of the experimental design were as follows:

1. The messages were first divided into three silos. A silo comprises messages of a similar type. The primary use of the silo is as a bookkeeping device. The silo itself has no other function than to ensure that only one message from a set of similar messages would appear in a test concept. An example of two message components that would be placed into the same silo is the following pair, which should not appear together: *Now there is a new toothpaste that bonds calcium to your teeth for protection when you need it most* and *Now there is a new toothpaste that has concentrated calcium that penetrates teeth.*

2. The appropriate messages were then put into their respective silos. Table **1** gives an example of these silos and six messages from each silo.

3. A total of 120 different concepts (or combinations of message components) were created, with the property that the components were statistically independent in these combinations. The independence allowed for subsequent analysis by regression modeling, in order to estimate the impact or contribution of each element. A stratagem of having incomplete test concepts was employed to avoid multicollinearity. The stratagem allowed the regression analysis to generate estimates of the absolute magnitude of contributions of the components, not relative contributions, which would be the case were multicollinearity to exist.

Table 1: Best Scoring and Worst Scoring (Italicized) Components for Three Countries, Along with the Additive Constant.

	Total	**Philippines**	**Australia**	**Mexico**
	367	120	127	120
Additive constant	76	75	63	38
Philippines				
Now there is a new toothpaste that has both liquid calcium and super-charged fluoride for a dual defense system	0	4	–1	2
So it provides protection from cavities superior to professional strength fluoride	2	4	3	3
If you are a parent of young children, you know their teeth are vulnerable and need maximum cavity protection so they can stay cavity-free	1	4	0	4
So it gives protection from cavities as good as a fluoride treatment from the dentist	4	4	**6**	3
Now there is a new toothpaste that replenishes the minerals lost from your teeth throughout the day	–2	–6	3	–2
Australia				
So it provides protection from both children's cavities and the cavities that adults get at the gum line and around existing cavities	3	0	**9**	2
Now there is a new toothpaste that bonds calcium to your teeth for protection when you need it most	1	–2	**9**	–10
Now there is a new toothpaste that has concentrated calcium that penetrates teeth	2	0	**8**	0
So it protects against cavities between meals better than other toothpaste	2	0	**8**	**6**
Now there is a new toothpaste that bonds calcium and fluoride to your teeth for protection when you need it most	2	–1	**8**	–3
So it rebuilds, restores and strengthens tooth enamel	3	1	**8**	–1
Now there is a new toothpaste that provides a mineral reservoir for your teeth	–4	0	–7	1
Mexico				
So it regenerates enamel twice as fast so cavities can't get started	3	3	4	**10**
So it actually repairs the early stages of cavities throughout the day, even while you're sleeping	2	1	4	**9**
So it provides cavity protection in hard-to-reach areas where your toothbrush doesn't reach	3	–4	**7**	**8**
Now there is a new toothpaste that provides calcium protection when and where your teeth need it most	2	1	**6**	–6
Now there is a new toothpaste that has a slow release calcium formula for sustained calcium protection	0	0	**7**	–10
Now there is a new toothpaste that bonds calcium to your teeth for protection when you need it most	1	–2	**9**	–10

Note. For the Philippines and Australia, interest was measured by the top three boxes on a 9-point scale. For Mexico, interest was measured by the top one box on a 9-point scale. Poor performing elements are shown in italics.

FIELD EXECUTION

The study was run with respondents pre-recruited in each market to participate in a central-location test. The respondents were recruited to participate for a two-and-a-half-hour test session and were paid at the end of the session. During the session, the respondent first read an introductory page about the purpose of the study (to evaluate new ideas for toothpastes) and then evaluated all 120 concepts in a randomized order. The order of test concepts was changed for each individual. The respondent rated each concept on an anchored 1–9 scale for interest. After completing the concept evaluations, the respondent completed a self-profiling classification.

In Australia and the Philippines, the concepts were presented in English. In Mexico, the entire study—instructions, as well as concepts—were presented in Spanish and evaluated using a Spanish rating scale. The entire study was translated into Spanish and then "back-translated" by a different service. The back-translation ensures that the translation parallels the English version.

Altogether 367 respondents participated: 127 in Australia and 120 each in Mexico and the Philippines. Fig. **1** represents the distribution of respondent ratings for the three country markets. As discussed in Appendix 1, an adjustment to some criteria in the analysis had to be made because of the exceptionally high positive responses to the concepts. Mexican respondents consistently up-rated the concepts, not necessarily because they liked the concepts more, but rather because of the well-known bias to up-rate test stimuli. That bias was well-known and confirmed by the field service in Mexico City. A more detailed discussion of the distribution of ratings and the choice of response criterion appears in Appendix 1.

CREATING THE CONCEPT–RESPONSE MODEL

Each of the respondents evaluated all 120 concepts in a randomized order. As the communication or message components were presented in a way that made them statistically independent, it was possible to create an individual level model for each respondent. The individual level model was created after a simple binary transformation was made to the ratings. For Australia and the Philippines, the ratings of 7–9 were converted to 100 and the ratings of 1–6 were converted to 0.

For Mexico, which showed much higher ratings, ratings of 9 were converted to 100 and ratings of 1–8 were converted to 0. This transformation follows the conventions of market research, which looks at incidence statistics (how many people fall into a specific group) rather than intensity statistics (how strong is the feeling among respondents in the group; Moskowitz, Porretta and Silcher, 2005).

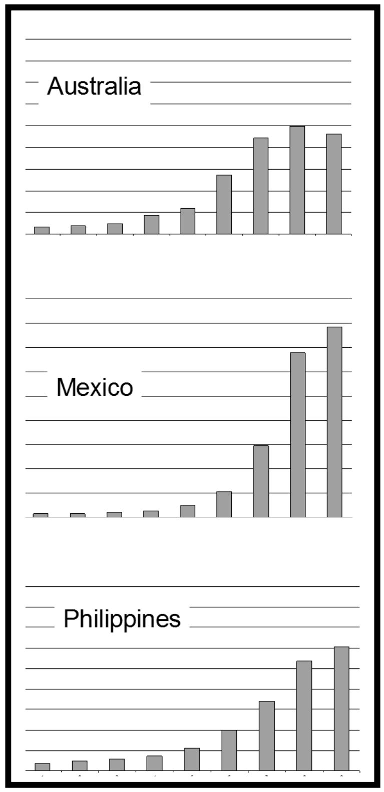


Figure 1: Distribution of ratings for the 120 concepts by all respondents in a country, on the 9-point scale. Starting at the left, the first bar shows the number of ratings for scale value 1. At the far right, the ninth bar shows the number of ratings for scale value 9. Australia shows 70% of concepts for the top three box, Mexico shows 88% and the Philippines shows 74%.

The additive model can be expressed by the simple linear equation:

Rating = $k_0 + k_1$ (Insight #1)… k_{13} (Insight #13) +

k_{14} (Reason to Believe #1)… k_{49} (Reason to Believe #36) +

k_{50} (Benefit #1)… k_{92} (Benefit #43)

The equation was estimated by the method of ordinary least-squares regression (Systat, 1997). The additive constant measures the predisposition of a respondent to be interested in the toothpaste concept if no elements are present in the concept. Thus the additive constant may be considered to be a baseline parameter, showing the proclivity of the respondents to accept the concept.

RESULTS

Table **1** reveals that the additive constant is high for the Philippines (75) and somewhat lower for Australia (63). For the Philippines, this means that without any specific elements present, 75% of the respondents would be interested in the concept. In contrast, with Australia, about 63% would be interested. For Mexico, only 38% would be interested. A brief discussion of the coefficients, *i.e.* the utility values of the components, appears in Appendix 2.

The analysis of results by country suggests country-to-country differences that cannot easily be overcome. On a practical basis, it may be necessary to change the rules. One change is to find winning elements that work in the three countries but among a defined subgroup of individuals, *e.g.,* segments. The strategy searches for commonalities among consumers, independent of country. By segmenting individuals worldwide, independent of country, one may discover homogeneous groups of consumers that reside in three different countries. With the homogeneity comes the possibility of developing a single product concept for toothpaste, which applies across all three country markets and which communicates the same type of selling points in a coherent way. The benefits are that this single product concept can be more targeted because the mind-set is homogeneous. The notion of country becomes one of "accident" to be dealt with by modifying the basic concept for the segment. Thus the concept can thus be transported worldwide. Homogeneity in this context is defined in terms of preferences or propensity to buy.

MIND-SET SEGMENTATION: RATIONALE AND MECHANICS

It should be noted that numerous methods have been proposed for segmenting consumer markets in international settings. These methods include aggregate and disaggregate country-level variables (Steenkamp and Ter Hofstede, 2001) and psychographic criteria (Ter Hofstede *et al.* 1999). The importance of addressing the possibility of behavioral heterogeneity and homogeneity within and among consumer cultures and the predictive validity of segment solutions has also been established (Wedel and Kamakura, 2000). However, engineering messages anchored in the behavioral homogeneity of segments, with these segments cutting across national boundaries, has not been well researched. More often than not, consumer decision making and behavioral homogeneity issues are evaluated from both an inter- and intra-*cultural* focus (Broderick *et al.* 2007).

In this study, we focus on the various components of a promotional communication. We further deal with the steps involved in understanding what to say to transnational segments in order to generate the most effective message for a particular segment. We identify well-defined segments from the perspective of the buyer's mind, using RDE to understand these segments.

The clustering of individuals into like-minded segments includes the following steps:

1. Each respondent evaluates all the test concepts.

2. The data from each respondent allow for a separate and complete model relating the 9-point rating the concepts to the presence/absence of the 92 concept elements. This is the *persuasion model*, which shows how many rating points are contributed by each element. The persuasion model does not involve the transformation into a binary, as discussed above. The only use for the persuasion model is to develop the coefficients or utilities to be used in the segmentation.

3. Based on the 92 utilities, one per element, one set per respondent, calculate "distance" between pairs of respondents using the statistic (1-*R*), where *R* = Pearson correlation coefficient. When the persuasion coefficients for two respondents correlate perfectly, so that increases in one parallel increases in the other, then the two respondents show identical patterns. The correlation, *R*, is 1.0 and the distance between them by the (1-*R*) statistic is 0 (viz., 1-1 = 0). When the persuasion coefficients for the two respondents inversely relate, so that increases

in one correspond to precise decreases in the other, then the correlation is –1 and the distance is therefore 2 (viz. 1- –1 = 2).

4. The cluster program assigns the respondents to homogeneous clusters or groups, such that the distance between pairs of respondents in a group is small, whereas the distance between the centroids of the group is large. The statistics for the method are known as k-means clustering, with the distance being defined by the Pearson correlation (SYSTAT, 1997).

5. The ideal solution is a minimal number of clusters (defined statistically) that can be easily interpreted (defined subjectively). Interpretation is as simple as the winning elements being coherent, making sense and allowing for a story. In these results, it appeared that the three-cluster solution was easiest to interpret.

The analysis revealed three segments, with similar distributions across the three markets (see Table **2**):

1. Segment 1—Superior technology (Tech)—wants new technology that is better than currently available.

2. Segment 2—Deliverable results (Results)—wants to know that it fights cavities and repairs teeth; benefit oriented.

3. Segment 3—Essential safety seekers (Safety)—wants to know that it works and also that it is safe.

4. The Tech segment has the highest basic interest (82) as shown by the additive constant, whereas the Safety segment has the lowest basic interest (70). The Tech segment is ready to buy the new product, presumably based on the orientation page of the study, which talked about technology breakthroughs.

5. The Results and Safety segments show a number of breakthrough elements having utility values (interest model) of +8 or higher.

6. Results of the experimental design used in this study suggest creative strategies for the product category selected can be standardized by segment in Australia, the Philippines and Mexico.

7. Different configurations of selling propositions can be put together to generate an effective advertising messages or even themes targeting the three identified segments. Similar results were derived by Wei and Jiang (2005) using a very different methodology (analyzing magazine ads for a cell phone in China and the USA).

Table 2: Distribution of Segments by Market and Performance of Winning Elements for Toothpaste Across the Three Mind-Set Segments.

	Total	**Tech**	**Results**	**Safety**
Base size (number of respondents)	367	155	122	90
Australia	128	53	41	34
Mexico	120	48	42	30
Philippines	119	54	39	26
Distribution of the three segments in each of the three countries				
Australia	100%	41%	32%	27%
Mexico	100%	40%	35%	25%
Philippines	100%	45%	33%	22%
Additive constant	76	82	75	70
Segment 1—Technology seekers				
Now there is a new toothpaste that has a calcium complex that adheres to and repairs weak spots	2	6	–6	5
Segment 2—Results oriented				
So it locks out cavities, locks in protection	3	–2	**10**	3
So it adds years of life to your teeth	3	–5	**9**	6
So it provides cavity protection in hard-to-reach areas where your toothbrush doesn't reach	3	–1	**8**	4
So it gives protection from cavities as good as a fluoride treatment from the dentist	4	2	**8**	0
So it stops cavities before they start	3	–1	**8**	3
Segment 3—Protection and tooth safety seekers				
Now there is a new toothpaste that has a slow release fluoride formula for sustained fluoride protection	1	–1	–4	**11**
So it provides protection from both children's cavities and the cavities that adults get at the gum line and around existing cavities	3	–1	2	**10**

Table 2: cont….

So it keeps teeth in the safety zone to keep them cavity-free	1	–5	4	**9**
Now there is a new toothpaste that has unique calcium-activated fluoride	2	1	–1	**8**
Now there is a new toothpaste that has concentrated calcium that penetrates teeth	2	2	0	**8**
Now there is a new toothpaste that has a calcium-fluoride-complex that adheres to and repairs weak spots	1	2	–4	**8**

Note. Strong performing elements are shown in bold numbers.

DISCUSSION

The debate surrounding standardization *versus* localization has raged ever since Levitt (1983) introduced the concept of "segment simultaneity". In an increasingly global and technology-savvy marketplace, customer segments are becoming homogenized across national boundaries. A consequence of this homogenization is that behavioral and lifestyle segmentation may be a necessary addition to geopolitical and economic segmentation in international markets (Wedel and Kamakura, 2002).

More recently, scholars examined the empirical relation between positioning strategies and segmentation in international markets (Hassan and Craft, 2005). The argument is that international segmentation should be linked to strategic positioning decisions in an increasingly competitive and transparent marketplace. Segmentation must serve the business goals in taking advantage of the already-existing brand equity.

A second benefit of international segmentation is improving knowledge. This benefit is the increased relevance of the knowledge workers. Models of messaging that simultaneously segment the international market and identify an optimal positioning intra- and/or internationally increase the managerial relevance of researcher devoted to market segmentation (Steenkamp and Ter Hofstede, 2001).

Segmentation can also achieve economies of scale by creating the same product appropriate to different countries.

Our data cast some light on another area of interest to marketers. Recently, experts and pundits have announced that there are opportunities to customize marketing by country, taking into account the mind-set of the customers. Customer purchase behaviors are diverse within and between nations (Gilmore and Pine, 1997). RDE as deployed in this study extends this stream of academic research by examining appropriate messaging strategy as related to international market segmentation in a relatively high-volume product category—toothpaste. Specifically, we present an approach that includes within and across country/market evaluations. The approach aids in structuring the homogeneity that may exist among consumers and nations using RDE-based segmentation, working as it does with the actual messages.

Market dynamics provide another area of business where RDE-based segmentation may contribute. Escalating media costs, increasing communication and linkages across markets and the internationalization of retailing (Douglas and Wind, 1987) force multinational companies to feature fewer brands in their international markets. The rationale for these fewer products is reduced costs. The segmentation scheme that emerges from the RDE study helps the multinational company design the appropriate positioning for its brands. The right communication selects creative elements that complement each other and that work in many countries. The company has less to do and the consumer has less to learn. Fewer brands, but stronger messaging, drive easier communication and a greater chance of consumer acceptance.

MESSAGING, RDE AND A VIEW OF THE FUTURE

In a world awash with choices, with competing brands and with rapidly changing technology, it is important to ensure that the messages one sends to consumers are important in terms of persuasion, as well as actionable in terms of technology. Consumers may not be able to design products "out of their minds," but certainly they can react to products. This study suggests that traditional bases for segmentation (*e.g.,* products used, countries, *etc.*) may not necessarily be as powerful as segmentation based upon *patterns* of responses.

There are three key implications for the future use of RDE in designing transnational communication based on segments:

1. **Scope:** It is important to test *many* different messages. The focus should be on the stimulus (messaging), not on the respondent (country).

2. **Trend searching for evolving mind-sets:** Look for trends by doing the experiment on a periodic basis (*e.g.,* yearly or every other year). Studies will search for changes in the segmentation over time, as a result of changing consumers and changing competitive frames.

3. **The "who" of the segment:** Furthermore, future research is called upon to learn more about the segments. What is common about the individuals who fall into a segment? The "why" is not necessary to create the segments, but it may lead to new insights about consumers. The segmentation, however, can still be used to target development and messaging, even if the "why" is unanswered.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Aaker, J. & Williams, P. (1998). Empathy *versus* pride: The influence of emotional appeals across cultures. *Journal of Consumer Research, 25*(3), 241–262.

Avery, J. & Keinan, A. (2008) Understanding Brands. Case Study 9-509-041. Boston, MA: Harvard Business School Publishing.

Barwise, P. & Farley, J. (2005). The state of interactive marketing in seven countries. *Journal of Interactive Marketing, 19*(1), 23–26.

Boddewyn, J. J., Soehl, R., & Piccard, J. (1986). Standardization in international marketing: "Is Ted Levitt in fact right?" *Business Horizon, 29,* 69–75.

Broderick, A. J., Greenley, G. E., & Dentiste Mueller, R. (2007). The behavioural homogeneity evaluation framework: Multi-level evaluations of consumer involvement in international segmentation. *Journal of International Business Studies, 38,* 746–763.

Czinkota, M. R. & Ronkainen, I. (2007). *International marketing* (8th ed.). New York: South-Western.

De Paola, D. P. (Ed.). (2007). *Colgate Care Report, 17*(3). Available at https://secure.colgateprofessional.com/app/cop/repository/article-310/frameset.jsp?middle=ocrindex.html (Retrieved on 05.12.2010)

Douglas, S. & Wind, Y. (1987). The myth of globalization. *Columbia Journal of World Business, 24*(4), 19–29.

Duncan, T. & Ramaprasad, J. (1995). Standardized multinational advertising: The influencing factors. *Journal of Advertising, 24,* 55–68.

Fastoso, F. & Whitelock, J. (2007). International advertising strategy: The standardisation questions in manager studies. *International Marketing Review, 24* (5), 591–605.

Gilmore, J. H. & Pine, B. J. (1997). The four faces of mass customization. *Harvard Business Review, 75*(1), 91–101.

Hassan, S. S. & Craft, S. H. (2005). Linking global market segmentation decisions with strategic positioning options. *Journal of Consumer Marketing, 22*(2), 81–89.

Hult, G. T., Tamer Cavusgil, S., Deligonul, S., Kiyak, T., & Lagerström, K. (2007). What drives performance in globally focused marketing organization? A three-country study. *Journal of International Marketing, 15*(2), 68–85.

Information Resources Inc. (IRI). (2007). Various documents, Information Resources Inc. Retrieved 04.12.2010 from http://us.infores.com/

Jain, S. (1989). Standardization of international marketing strategy. *Journal of Marketing, 53*(1), 70–79.

Keller, K. L. (2001). Mastering the marketing mix: Micro and macro perspectives. *Journal of Marketing Management, 17*(29), 819–847.

Ko, E., Kim, E., Taylor, C., Kyung, H. K., & Kang, J. J. (2007). Cross-national market segmentation in the fashion industry. *International Marketing Review, 24*(5), 629–651.

Levitt, K. (1983). The triad world view. *Journal of Business Strategy,7,* 8–19.

Levitt, T. (1993). The globalization of markets. *Harvard Business Review, 61*(3), 92–102.

Moskowitz, H. R., Porretta, S., & Silcher, M. (2005). *Concept research in food product design and development*. Ames, IA: Blackwell.

Orth, U. R. & Firbasova, Z. (2003). The role of consumer ethnocentrism in food product evaluation. *Agribusiness, 19*(2), 137–153.

Papavassiliou, N. & Stathakopoulos, V. (1997). Standardization *vs.* adaptation of international advertising strategies: Toward a framework. *Journal of Marketing, 31,* 504–527.

Rigby, D. K. & Vishwanath, V. (2006). Localization—The revolution in consumer markets. *Harvard Business Review, 84*(4), 82–92.

Stafford, M. R. (2005). International services advertising. *Journal of Advertising, 34*(1), 65–86.

Steenkamp, J. B. E. M. & Ter Hofstede, F. (2001). International market segmentation: Issues and perspectives. *International Journal of Research in Marketing, 19*(3), 185–213.

SYSTAT (1997). *The system for statistics. User manual.* Evanston, IL: Systat Corporation, Division of SPSS.

Szymanski, D. M., Bharadwaj, S. G., & Varadarajan, P. R. (1993). Standardization *vs.* adaptation of international marketing strategy: An empirical investigation. *Journal of Marketing, 57,* 1–7.

Ter Hofstede, F., Steenkamp, J.-B. E. M., & Wedel, M. (1999). International market segmentation based on consumer-product relations. *Journal of Marketing Research, XXXVI,* 1–17.

Theodosiou, M. & Leonidou, L. C. (2003). Standardization *vs.* adaptation of international marketing strategy. *International Business Review, 12*(2), 141–171.

Wedel, M. & Kamakura, W. A. (2000). *Market segmentation: Conceptual and methodological foundations*. Dordrecht, the Netherlands: Kluwer Academic Publishers.

Wedel, M. & Kamakura, W. (2002). Introduction to the special issue on market segmentation. *International Journal of Research in Marketing, 19*(3), 181–183.

Wei, R. & Jiang, J. (2005). Exploring culture's influence on standardization dynamics of creative strategy and execution in international advertising. *Journal of Mass Communication Quarterly, 82*(4), 838–856.

Whitelock, J. M. (1989). Global marketing and the case for international product standardization. *European Business Review, 89*(3), 32–44.

APPENDIX 1

Distribution of ratings and the choice of response criterion.

Traditionally, consumer researchers divide the response scale into two or more sections and then count the proportion of responses in each section of the scale for each concept. In the USA and Western Europe, where many of these studies are run with the 9-point scale, the convention has been to select ratings of 7–9 to denote interest in the concept and ratings of 1–6 to denote lack of interest. This convention separates winning from losing ideas.

With respondents in other cultures, however, sometimes the choice of the 1–6 and 7–9 generates problems. It does in Mexico, where the tendency is to up-rate most concepts, thus giving a false positive. We saw this evidence for false positives in Fig. **1**, which shows that for Mexico, 88% of the ratings are 7–9. Thus for Mexico we had to change the criteria of what we accept as a positive reaction, as the overwhelming majority of ratings are such high positives. For this study, as well as for others, we looked at increasing the criteria from 7–9, first to 8–9 and then finally to 9 alone. We found that only when we used the top single rating, 9 alone, did we begin to see real differences in the reactions to the concept. Thus, we changed the rules of the convention to align with the biased use of the scale by Mexican respondents, counting only a rating of 9 (rather than 7–9) as a positive vote.

APPENDIX 2

The initial results for the additive constant raise an important issue for international studies of the algebra of the consumer mind. Marketers and researchers like to look at numbers that are commensurate with each other across countries. Thus the ideal situation occurs when one can compare the ratings in one country with the ratings in another country. When two countries have top three boxes of 80 *versus* 40 and respondents in the two countries use the scale similarly, we can conclude that twice as many respondents in the first country are interested in the basic proposition *versus* those in the second country. In contrast, when the respondents use the scale differently, as we suspect in Mexico, which up-rates most of the concepts (88% use the ratings 7–9), then we cannot validly draw that conclusion due to the scaling bias. We are left with the quandary—either use the same measure for concept acceptance with little differentiation among messages by Mexican respondents, or impose different rules on analyzing data, with these rules varying by country and the criteria for acceptance made more stringent in some cases, such as that of Mexico. We followed the latter approach here, with the consequence that we have lower additive constants for one country (Mexico) than for the others, because we have made the criteria for acceptance more stringent in light of scaling behavior.

The key information for toothpaste marketers lies in the utility values of the components, *i.e.* the coefficients of the equation above. Each component generates its own coefficient, k_i, which can be interpreted as the additive (or subtractive) percent of respondents interested in the toothpaste. The utility values can be added together with the additive constant to generate an expected percent of the respondent population interested in the toothpaste. (Keep in mind that for Mexico this means the percent of respondents who rate the concept 9; for Australia and Philippines this means the percent of respondents who rate the concept 7, 8, or 9.) At the end of the day we will either have a higher additive constant and lower utility values, or a lower additive constant and higher utility values, respectively.

CHAPTER 12

Extending RDE to Evaluating Potential Social Anxiety Factors

Howard R. Moskowitz[1,*], Alex Gofman[1], Gillie Gabay[2], Hollis Ashman[3] and Ipshita Ray[4]

[1]Moskowitz Jacobs Inc., White Plains, New York, USA; [2]College of Management Academic Studies, School of Behavioral Management, Hamerkaz, Israel; [3]Muse LLC, Powell, CO and [4]Pace University, Lubin School of Business, New York, New York, USA

Abstract: This chapter addresses the question of how to create a database of the citizen's mind about anxiety-provoking situations in the face of terrorism. The approach is grounded in a combination of experimental design, psychophysics (a branch of psychology) and consumer research. The theoretical foundation is illustrated with a set of 15 empirical studies using conjoint analysis in order to understand how consumers respond to anxiety-provoking situations. The approach identifies the mind-set toward terrorism at the individual respondent level. By exploring responses embedded in a general study of anxiety-provoking situations, it becomes possible to understand the algebra of the individual respondent's mind, how important the basic fear of terrorism actually is, how important it is to specify the type of terrorism (bombing *versus* contamination of the food supply) and the structure of what frightens the consumer. The chapter attempts to answer the question: what are the critical drivers of anxiety—the specific terrorist act, the location of the act, the feelings, or even the proposed remedies to reduce anxiety? The outcome of the research is both an empirical dataset and potentially a framework for a subdiscipline in social science. This approach looks at problems from three perspectives: as a scientist—to understand general patterns; as an engineer—to solve a specific problem; and as a clinical psychologist—at the level of a single individual (idiographic) as well as at the level of the general population (nomothetic).

Keywords: Social anxiety, rule developing experimentation, terrorism, reducing anxiety.

INTRODUCTION

Social marketing communications frequently use scare tactics or appeals to fear in order to convince people to adopt desired alternatives (Albrecht, 1996; Donovan and Henley, 1997). Other scholars argue that people just need to be prepared for

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

terrorism (Marshall *et al.* 2007). Although both approaches acknowledge the reality of fear and terror in the population, neither approach addresses the issue of finding appropriate public communications to calm citizens' angst related to the topic.

People are vitally interested in the way their government treats them and what the government does in its role as protector of the people (Annas, 2002). The topics related to anxiety about terrorism and the search for meaningful measures that allay this anxiety, have been frequently researched in the recent literature (Annas, 2002; Smelser and Mitchell, 2002; Eisenman, *et al.* 2009). This chapter deals with a psychological effect of terrorism.

According to Tanielian and Stein (2005), there is an urgent need to develop effective tools to measure the impact of psychological, social and political responses. The threats vary, from the impact of an attack or threat and its consequences, to the impact of personal preparedness by each citizen (*e.g.,* behavioral and social procedures), or acts of counterterrorism by governments (*e.g.,* new security institutions). Authorities around the world have undertaken unprecedented efforts to increase the nation's ability to respond to terrorism. However, little has been done to focus on the importance of addressing the *emotional* consequences of terrorism as part of counterterrorism.

In the course of their review of databases on terrorism, Lum, Kennedy and Sherley (2006) discovered that there is an almost complete absence of evaluation research on these types of counterterrorism strategies. They conclude that there is little scientific knowledge about the effectiveness of most counterterrorism interventions. Their literature review revealed that some interventions either did not work or occasionally actually increased the likelihood of terrorism and terrorism-related anxiety. The findings of this review dramatically emphasize the need for government leaders, policy makers, researchers and funding agencies to include evaluations of the effectiveness of these counterterrorism programs in their agendas.

Qualified researchers need to be included in counterterrorism policy making, strategic thinking, planning and evaluation. More of the research on terrorism and counterterrorism needs to be empirical and evaluative, using scientific principles and different types of methodology. This chapter describes and quantifies the psychological consequences of terrorism and outlines response strategies for dealing with them. Such information should prove useful for policy makers attempting to develop state and local response strategies.

STIMULUS–RESPONSE METHODS AND PSYCHOPHYSICAL THINKING

Despite the popularity of questionnaires, they are limited to data collected from responses to the questions that are asked. Furthermore, there is the ever-present respondent tendency to please the interviewer in a personal interview, or to be "politically correct" either in the personal interview or in a paper- or computer-based interview. The tendency to comply with interviewer demands is well known in the psychological and social sciences literature (Rosenthal, 1976) and can lead to biased results, even when the respondent is not aware of such biases generated by the interview.

A different way to look at issues of social policy can be implemented using methods originating from physics and chemistry but adopted by experimental psychologists. These methods go by the rubric of stimulus–response methods and trace their origin to the philosophy of *operationism* wherein knowledge is defined as the ability to "effect" a specific action by knowing what aspects of the antecedent conditions to change. The guiding principles are based upon the logic of experimentation. For social scientists the ingoing belief is that the key learning comes from the pattern of responses to test stimuli. When the respondent is presented with a series of test stimuli and the ratings to these stimuli are obtained, the relation between what the researcher presented and how the respondent scored the test stimuli provides the key information. The respondent need not even be aware of the criteria underlying his scoring. The regularity of such patterns and the ability to uncover the underlying relations between variables and responses is what constitutes the science.

The application of stimulus–response thinking to the world of social science can be traced back a century and a half to the seminal thinking of psychophysics, the first branch of experimental psychology and the inspiration for the approaches discussed here. Psychophysics searches for orderly relations between what we perceive through our senses and the nature of the physical stimulus, usually the magnitude but often the quality of that stimulus. The goal of psychophysics is to develop relations between variables. It is these relations that generate the substance of our knowledge about how we perceive stimuli and transform those stimuli into subjective responses. Psychophysics is informed by physics and chemistry, especially by the search for "rules" or at least for "regularities" in nature (Stevens, 1975).

HOW DOES PSYCHOPHYSICS FIT WITH SOCIAL POLICY?

It is easy to trace the history of psychophysical thinking on the senses and to identify how it evolved out of simple testing of differences to developing equations, which show the change in perceptual intensity *versus* physical magnitude (*e.g.,* changes in sweetness with increasing amounts of sugar). The psychophysicist studying the private world of sensory perception needed simply to change the amount of sugar in a water solution, creating thereby a number of test stimuli, present these solutions in some randomized order to a respondent, get a rating of "sweetness" or "liking" (depending on the specific issues being addressed) and then plot the rating against the physical level of sugar.

But what about social policy, where there is no intrinsic metric as we have for sugar? At its very basic level, psychophysics can be viewed as an application of experimental design as statisticians conceive of such a discipline (Box, Hunter and Hunter, 1978). The psychophysical way of thinking conceives of variables in experimental design as physical stimuli that are mixed and matched. Psychophysics thinks, in turn, of the respondent as the device that integrates this information about the mixtures and comes up with a response, which is deconstructed by statistical analysis into the contribution of the individual components.

Following this train of thought, let us move forward in our work on social policy using psychophysical thinking and experimental design. We will treat social issues as simple, stand-alone phrases that can be mixed, matched and presented to respondents to obtain ratings. We have specific independent variables (the phrases) and a measured response (*e.g.,* anxiety). Rather than seeing how the response alerts with changes in one variable, we measure the contribution of each of the individual variables to the response, with these variables being statements that are either present or absent in the test stimulus. We have recreated a psychophysical design, albeit with the stimulus taking on the value 1 when present, or 0 when absent.

The application of psychophysical thinking and experimental design begins with the method of conjoint analysis. The objective of conjoint analysis is to understand how components of mixtures act from responses to the mixture. Early conjoint analysis approaches to public policy were demonstrated in Moskowitz, Gofman, Tungaturthy, Manchaiah and Cohen (2000).

The approach described in Moskowitz and Gofman (2007) first identifies the raw materials to be studied, which in the case of public policy comprises relatively single-minded, stand-alone phrases dealing with the different facets of a social issue. These phrases are classified as belonging to different silos, or categories. The silo or category thus comprises like-minded elements or ideas, which may differ in what they convey. The elements are mixed and matched by experimental design to create combinations. The elements appeared independently of each other in a statistical sense, although it is hard for a respondent to discern the underlying design. The respondent rates the combination, *i.e.* the test concept, on a scale. The ratings are then analyzed to show the number of scale points contributed by each component.

THE PSYCHOPHYSICS OF TERRORISM-BASED ANXIETY

The terrorism study, conducted in 2003, was one of 15 different studies run as part of the Deal With It!®™ database. Each of the 15 studies was constructed in the same way: four major silos, each with nine elements. (Note: These studies were performed by It!® Ventures LLC, a collaboration between Moskowitz Jacobs, Inc. and The Understanding and Insight Group, Inc.).

Studies dealing with anxiety are in some ways intrinsically frightening because they address issues that are unpleasant. Unlike traditional consumer research studies dealing with food, with shopping and the like, studies in the Deal with It!® database were clearly addressing the topics that many people would rather forget. We can get a sense of the relative "anxiety" produced by these topics ahead of the actual study through completion rate analysis of how many respondents logged in to participate in the study and how many of those actually completed. The typical completion rate is between 50% and 70% for studies dealing with more pleasant, non-anxiety-provoking topics such as shopping. This project showed lower completion rates—from 31% (health-care system) to 60% (loss of assets).

The heart of the study is the set of different experimentally designed concepts, which comprise the distinct elements from the study, mixed and matched to create vignettes. The elements for terrorism were selected to range from relatively light to severe. We see the range in Table **1**, which also contains the utilities of the elements for the total panel and subgroups. The utility values were estimated from regression after the rating scale was transformed from the original 1–9 anchored scale to a 0–100 anchored scale by a simple linear (affine) transform. We should look at elements with utility values above +15 as extremely strong drivers of anxiety; utility values above +10 as strong drivers; utility values above +5 as drivers; and utility values 0–5 as irrelevant. We should look at all elements with utility values below 0 as *reducers of anxiety*.

Table 1: Utility Values for the 36 Elements by Total Panel, Gender and Four Age Groups.

	Total	**Gender**		**Age**			
		Male	**Female**	**31–40**	**41–50**	**51–60**	**61–75**
Base size	121	28	93	21	42	37	16
Additive constant	44	38	46	37	47	44	45
Silo #1—Threats							
A3 A bomb under your car…	**15**	**14**	**16**	**21**	**16**	**12**	**11**
A9 A dirty nuclear bomb set off…	**15**	**15**	**16**	**22**	**15**	**12**	**10**
A4 Bombs blowing up in the middle of a building…	**12**	7	**13**	**18**	**12**	8	5
A7 A deadly disease like smallpox or anthrax let loose.	**10**	7	**11**	**17**	**13**	5	3
A6 Contamination of the food supply…	9	6	**10**	**16**	**10**	5	5
A5 Fire raging through a building…	6	1	7	**12**	5	4	0
A2 A bomb threat for a building that is a false alarm…	1	0	2	7	–1	–1	–1
A1 The media talking about potential terrorism acts…	0	–1	1	3	–2	0	2
A8 A computer virus let loose that impacts your everyday businesses…	–2	–1	–2	1	–4	–2	1
Silo #2—Location and target of the terrorism							
B3 An area crowded with children…	3	3	3	2	5	2	1
B9 During a red alert…	3	3	3	5	3	2	4

Table 1: cont….

B2 In a heavily populated area…	2	2	2	1	3	1	1
B5 An area filled with tourists…	2	0	2	2	3	–1	3
B6 You never expected it to happen to you or someone close to you.	2	2	2	2	1	2	1
B4 An area crowded with senior citizens…	1	3	0	1	1	3	–3
B7 During a yellow alert…	1	–4	2	1	1	–1	0
B8 During an orange alert…	1	1	1	2	1	–2	2
B1 In a nonpopulated area…	–2	–4	–1	–1	–2	–3	–1
Silo #3—How you respond to the threat							
C6 All the stress just builds up…you feel overwhelmed	3	1	3	5	3	3	0
C7 You experience temporary memory loss because there's just too much to take in.	2	3	2	7	0	5	3
C2 When you think about it, you just can't stop.	2	–1	3	3	2	2	4
C5 You experience it in all your senses…	2	0	3	5	1	3	***–6***
C4 You are scared…inside and out	1	4	1	3	1	2	–4
C1 You think about it when you are all alone…and you feel so helpless	1	–1	2	8	0	–1	–1
C9 At a turning point in your life.	1	1	1	3	0	1	–3
C8 Family and friends play a big role in your life…	0	1	0	0	–1	1	–3
C3 You'd drive any distance to get away from it…	0	–3	1	1	–1	1	***–5***
Silo #4—What might relieve the anxiety							
D2 You believe that international cooperation in the United Nations will keep you safe	**13**	**21**	**11**	8	**11**	**19**	**16**
D3 You think United Nations Forces will keep you safe	**12**	**21**	**10**	8	9	**19**	**13**
D5 You believe that the Centers for Disease Control will keep you safe	8	**10**	7	6	6	**10**	**11**
D4 You believe that Homeland Defense will keep you safe	7	**10**	6	5	7	8	**11**
D7 You think that your local hospital will keep you safe	6	7	5	2	6	6	**11**
D6 You think that your local police will keep you safe	6	7	5	6	5	4	**8**
D8 The media will keep you informed	–3	–1	–3	0	–2	–3	***–10***
D9 You need to contact your friends and family to make sure they are OK…	***–6***	–5	–7	***–13***	–3	***–9***	***–5***
D1 You trust that God will keep you safe	***–7***	4	***–10***	***–15***	***–9***	–1	–4

Note. Strong performing elements with utilities of 10 or greater are shown in bold; strong negative elements of –5 or less are shown in bold italics.

A conventional approach to studying responses to terrorism has the respondent rate each of these different elements on a scale, *e.g.,* degree to which this statement makes you "anxious". Such an approach is reasonable within a specific category or silo. What about the rating of elements in the fourth silo; those elements that are presumed to relieve the anxiety? Should these elements be rated as well on "anxiety," or do they need a different scale—the degree to which the anxiety is relieved? Furthermore, does this mean that we have to use two scales, or can we use a single scale, anchored at one end with a phrase such as “extremely anxious,” and at the other end with the opposite phrase such as “not at all anxious”? It is hard to develop a single scale that applies across the 36 elements, coming as they do from four different silos, with different meanings, different intents and different feelings attached to them.

Another way to think about these elements is to assume that they constitute building blocks that appear together in a single vignette. The vignette might comprise two, three, or even all four elements. Gofman and Moskowitz (2010) describe the details of the experimental design behind the approach. The vignette or test concept comprises one element from each of the four silos, but other vignettes might comprise fewer elements. The vignette represents the description of a situation. The respondent reads the vignettes and rates the entire combination on a 9-point scale such as ability to deal with the terrorism issue *(1 = Easily deal with it…9 = Cannot deal with it at all)*. This single response to a vignette is an easier task than rating the different, single elements, because people are accustomed to reading compound vignettes as completes, wholes, or so-called gestalts.

The actual analysis of the data is done on a respondent-by-respondent basis. Each respondent was presented with a unique set of 60 test concepts. The concepts were designed so that each element in a silo appeared independently of every other element in the other three silos. However, only one element from a silo (or no element from the silo) could appear in any one concept. This strategy ensured the elements from different silos were statistically independent, that across all of the respondents, allowing for analysis using ordinary least-squares regression. The provision that a concept need not have any element from another concept further ensured that the coefficients or utilities from the regression would have meaning in absolute terms. There was no collinearity, as often happens when the study is designed so that a concept must have one and only element from a silo.

ANALYSES—HOW AND WHAT DO THE DATA REVEAL USING EXPERIMENTAL DESIGN AND SELF-PROFILING?

More than two-thirds of the respondents appear to have free-floating anxiety regarding terrorism. We see this fact by looking at the distribution of the additive constant in Fig. **1** across the 121 respondents. Although the respondents were not directly asked about such free-floating, low-visibility anxiety, the additive constant shows the level of anxiety after the effects of the 36 individual elements have been partialled out. What becomes clear, however, is that the degree of such free-floating anxiety varies across people when we assume that the additive constant is free from biases due to scale usage. Of course, we cannot absolutely be certain of that freedom from bias, but the data do suggest a broad distribution, with the modal value near 0.

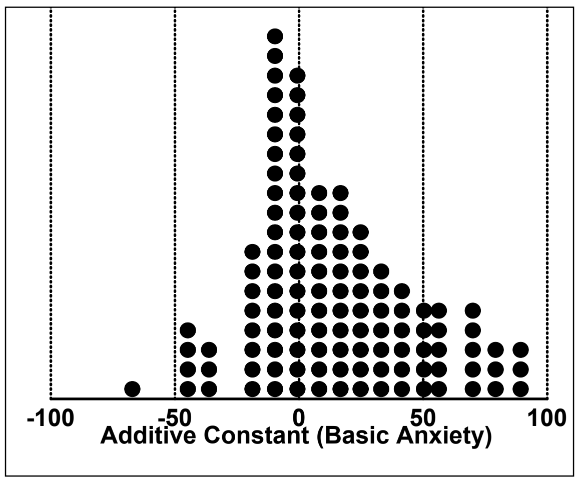


Figure 1: Distribution of the additive or baseline constant for the respondents. The abscissa has been truncated to the range from –100 to +100, so that some of the respondents lying slightly outside the range do not appear in this distribution. Baselines above 0 suggest a free-floating anxiety about terrorism; baselines above 50 suggest a much stronger free-floating anxiety; and baselines lower than 0 suggest little free-floating anxiety.

The analysis was been conducted for total panel and for different subgroups. Rule developing experimentation (RDE) generates a rich dataset, as we see from Table **1**.

1. The base sizes are different. There are about three times as many women as men participating in the study. This ratio is consistent with other RDE studies of this type, conducted on the Internet, with participation open to whomever wants to participate, which the authors have found consistent with other studies of this type. Researchers have found that for many studies it is easier to recruit women than it is to recruit men. In fact, it is necessary to put a

"screener" into place, when researchers want to balance the ratio of men and women. The screener ensures that an equal number of men and women participate. Once the quota for women is filled, the screener prevents any additional women from participating.

2. There is an inverted U-shaped curve for the base sizes, looking at age *versus* frequency. The greatest number of respondents fall into the group 41–60 years old, with correspondingly lower numbers of respondents falling into the groups 31–40 or 61–75, respectively.

3. The additive constant differs by gender and by age. In this RDE study, the additive constant shows the conditional probability of a person saying, "I cannot deal with the situation". The higher the additive constant, the greater the proportion of anxious respondents. Men are slightly less anxious than women (constant of 38 *versus* 46, respectively). Younger respondents (ages 31–40) are slightly less anxious than the older respondents (ages 41+).

4. Most of the positive utilities occur in silo #1, the threats. However, the threats are not all equally anxiety provoking. The most threatening are A3 ("A bomb under your car") and A9 ("A dirty nuclear bomb set off"). Both increase anxiety by (+15) points. These two threats produce anxiety among all of the different groups.

5. Most of the threats produce more anxiety among the younger respondents than they do among the older respondents. However, the differences by age are not always the same; *i.e.* it is not that older respondents are equally less anxious about all the threats. We see threats such as "Contaminated food supply" (A6) or "A fire raging through a building" (A5) being far more threatening to younger respondents than to older respondents. Quite possibly, some of these threats are seen as less probable by the older respondents than by the younger respondents, whereas all ages are exposed to news about bombings every day.

6. Computer viruses as threats are virtually irrelevant to these respondents, at least as a statement of a threat.

7. Silo #2 (location and target of the terrorism) is virtually irrelevant for respondents. Location and those affected may probably be informational, but that is all.

8. Silo #3 (how you respond to the threat) is also virtually irrelevant.

9. Silo #4 (what might relieve the anxiety) is the most surprising. Taken by themselves, these elements appeared to be reason anxiety reducers. For example, in the USA, the Centers for Disease Control is a well-known government body. Yet, merely mentioning this as an element in the vignette generates a lot of anxiety. In fact, it generates substantial anxiety among the total panel, among men and among older respondents (age 51+). This is quite unexpected. One might have thought that mentioning a government body would reduce anxiety, not increase it.

10. The same deleterious effect of government bodies appears when the vignette talks about the United Nations.

11. It is primarily trust in God and in friends that reduces anxiety. Furthermore, trust in God and in friends works more strongly to reduce the anxiety as experienced by the younger respondent than by the older respondent.

12. Trust in God actually increases a man's anxiety, whereas it strongly decreases a woman's anxiety (utility value of +4 for men; utility of –10 for women).

BEYOND THE TOTAL PANEL AND STANDARD GROUPING OF PEOPLE, ON TO MIND-SETS

Beneath the average, there must be different cross-currents, sometimes acting together, sometimes acting separately. Consumer market researchers and public opinion researchers know quite well that there exists intrinsic, often intractable variation among people so that the average may or may not represent the different individuals.

The conventional methods for dividing consumers create such groupings based on easy-to-acquire “exogenous information” such as demographics or self-profiled behaviors or attitudes. There certainly are some differences as seen from the gender scatterplot (left panel in Fig. **2**), but these differences are rather small.

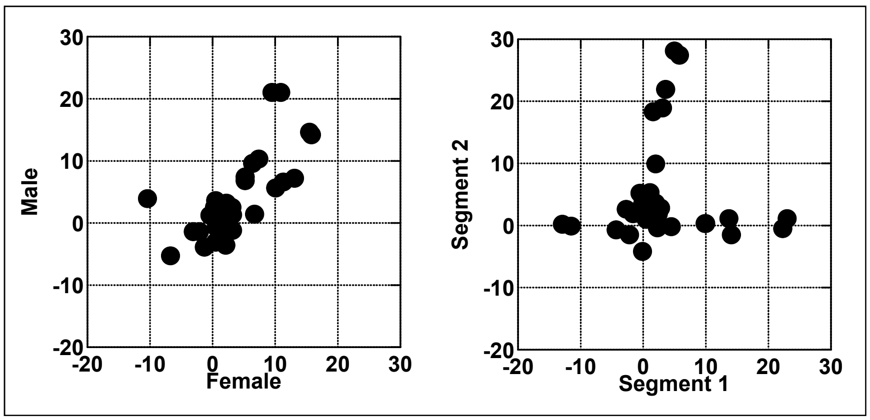


Figure 2: Scatterplots of the 36 utilities emerge from the study of terrorism. Each filled circle is an element. The left panel shows the scatterplot created according to gender. The right panel shows the scatterplot created according to mind-set clustering and segmentation. The respondents were divided into two groups whose patterns of utilities are most different from each other.

There are possible organizing principles that emerge when looking at the responses based on the place of living (Table **2**). The average utility for the first silo (threats) is much higher for individuals who live in an urban area (+11) than it is for individuals who live in a rural area (+6). Individuals in the urban area are simply more anxious. Conversely, when we talk about remedies, it is the rural respondents who are most concerned about government intervention. Urban individuals are not particularly responsive to government intervention. Finally, there are two actions or beliefs that are felt to keep one safe and to reduce anxiety:

1. The urban respondents feel much better when the message is about family. Specifically: “You need to contact your friends and family to make sure they are OK…” This is a strong anxiety reliever among the urban respondents (–10), but virtually irrelevant among rural respondents (–1).

2. The urban respondents feel indifferent when the utility talks about God, specifically: "You trust that God will keep you safe". This is irrelevant among urban respondents (–3), but a very strong anxiety reliever among rural respondents (–13).

Table 2: Average Utility of Elements by Nature of the Area in Which a Respondent Lives and the Utility Values for Individual Elements of Silo 4.

Text	City/Urban	City/Suburban	Midsize Suburban	Small Suburban	Rural
Mean by Silo					
Silo #1—Threats	**11**	9	8	6	6
Silo #2—Location and target of terrorism	4	1	0	1	2
Silo #3—How you respond to the threat	3	3	1	1	1
Silo #4—What might relieve the anxiety	1	5	2	2	9
Utility Values of Elements in Silo #4 (Presumed Anxiety Relievers)…Messages that Actually Relieve Stated Anxiety					
D9 You need to contact your friends and family to make sure they are OK…	***–10***	***–11***	–8	–6	–1
D1 You trust that God will keep you safe	–3	–4	–1	***–11***	***–13***
Utility Values of Elements in Silo #4 (Presumed Anxiety Relievers)…Messages that Actually Increase Stated Anxiety					
D2 You believe that international cooperation in the United Nations will keep you safe	2	**15**	7	**13**	**22**
D3 You think United Nations Forces will keep you safe	5	**15**	6	**10**	**22**
D4 You believe that Homeland Defense will keep you safe	3	6	6	6	**15**
D5 You believe that the Centers for Disease Control will keep you safe	2	**13**	8	5	**15**
D6 You think that your local police will keep you safe	2	8	1	4	**13**
D7 You think that your local hospital will keep you safe	2	4	6	4	**13**

Note. Strong performing elements with utilities of 10 or greater are shown in bold; strong negative elements of –5 or less are shown in bold italics.

Market researchers frequently cluster respondents by the pattern of the utilities, also known as concept–response segmentation (Moskowitz, Porretta and Silcher, 2005). The best way to find the description of these clusters sorts the elements by utilities in order to discover which specific elements perform most strongly each segment (see Table **3**).

1. The right-hand scatterplot in Fig. **2** suggests that the two segments differ dramatically. Respondents in Segment 1 look like they mostly afraid of external intervention. That is surprising when what one might have thought would reduce the anxiety of this segment actually enhances the anxiety.

2. Segment 2 is more like what we might think of as normal individuals. The segment is most anxious about actual terrorist actions, far more than one might have thought from the results generated by the total panel. Thus, these results suggest that there are at least two mind-sets in the population: those afraid of terrorist acts and those afraid of government relief. The latter group is unexpected, although their existence is not counterintuitive.

Table 3: Strongest Performing (Most-Anxiety-Producing) Elements for Total Panel and for the Two Concept–Response (*i.e.* Mind-Set Segments).

EL Text	**Tot**	**Seg 1**	**Seg 2**
Strongest elements for total panel			
A3 A bomb under your car…	**15**	6	**27**
A9 A dirty nuclear bomb set off…	**15**	5	**28**
D2 You believe that international cooperation in the United Nations will keep you safe	**13**	**23**	1
D3 You think United Nations Forces will keep you safe	**12**	**22**	–1
A4 Bombs blowing up in the middle of a building…	**12**	4	**22**
A7 A deadly disease like smallpox or anthrax let loose.	**10**	3	**19**
D1 You trust that God will keep you safe	*–7*	*–13*	*0*
Strongest elements for Segment 1—Anxiety from outside contact with a government agency (national or international)			
D2 You believe that international cooperation in the United Nations will keep you safe	**13**	**23**	1
D3 You think United Nations Forces will keep you safe	**12**	**22**	–1

Table 3: cont….

D4	You believe that Homeland Defense will keep you safe	7	**14**	–2
D5	You believe that the Centers for Disease Control will keep you safe	8	**14**	1
D7	You think that your local hospital will keep you safe	6	**10**	0
D9	*You need to contact your friends and family to make sure they are OK…*	*–6*	***–12***	*0*
D1	*You trust that God will keep you safe*	*–7*	***–13***	*0*
Strongest elements for Segment 2—Anxiety from actual terrorist acts				
A9	A dirty nuclear bomb set off…	**15**	5	**28**
A3	A bomb under your car…	**15**	6	**27**
A4	Bombs blowing up in the middle of a building…	**12**	4	**22**
A7	A deadly disease like smallpox or anthrax let loose.	**10**	3	**19**
A6	Contamination of the food supply…	9	2	**18**
A8	*A computer virus let loose that impacts your everyday businesses…*	–2	*0*	–4

Note. The most-anxiety-producing elements for each group ("agitating messages") are shown in bold. The least-anxiety-producing elements for each group ("calming messages") are shown in bold italics.

LOOKING AT INDIVIDUALS—THE POWER OF INDIVIDUAL-LEVEL MODELING

The analysis suggests is a hierarchy of terrorist incidents in terms of anxiety and a relatively poorly defined set of actions that a government can do in order to reduce the anxiety. *Indeed, when presented in vignettes, many of these so-called "remedies" to reduce anxiety in fact increase anxiety.* Respondents, not knowing that they should feel less anxious, actually say that the inclusion of these remedies make them feel even more anxious.

The general patterns of the utilities divide the respondents into different groups with different mind-sets. However, these general patterns do not give us a sense of how potential terrorist actions can be counteracted by specific government measures. Analyzing individual-level data allows us to see which individuals are sensitive to specific terrorist actions and to what specific communications, if any, these individuals respond.

As the regression modeling was done at the individual level, we can now look at the individual data in the following way:

1. The second silo (*where* terrorism occurs/among whom) and third silo (*response to the threat*) are both irrelevant. Respondents did not react

strongly to elements in either silo. We will further analyze only the first silo (terror incidents) and at the fourth silo (presumed remedies).

2. Classify each person as "anxiety prone" for a specific terrorism incident (silo #1) when the utility for the incident for that respondent exceeds a certain low value. Empirically, we choose the utility of +10, which means that the respondent says he is at least slightly more anxious than his baseline if this terrorism incident is present in the test concept. Any other cutoff can do as well; the +10 is simply an arbitrary threshold. When the person shows a utility >+10 for that terrorism incident, then classify the person as "1," *i.e.* the person is anxious. When the person shows a utility <+10 for that terrorism incident, then classify the person as "0," denoting the fact that the incident is not anxiety provoking.

3. Step 2 generates a new matrix of 1's and 0's for each person, for the nine terrorism incidents. In fact, a person can be sensitive to some incidents and not others.

4. Do the recoding of data, but this time focus on the presumed "*remedies*" listed in silo #4. However, the rules have to change for the recoding. We now look for those elements with utilities less than –10, which mean that the presence of the element in a concept reduces anxiety. We recode all utilities for this silo across respondents with a value less than –10 are recoded as "1" to denote them as anxiety relievers and the remaining utilities that are greater than –10 are recoded as 0.

5. Studying this new dataset, let's look at the correlation between the different terrorism incidents and their remedies. Are there any combinations where the terrorism incident increases anxiety and the remedy decreases anxiety *for that specific incident*? We correlate nine terrorism incidents, coded 1 or 0, with nine remedies, coded 1 or 0 using the appropriate correlation statistic for "binary data".

6. When we look at the total panel many of the correlations are quite low, which makes sense since the respondents fall into two clear segments.

7. Segment 1 shows very low correlations, near 0, because they are not as responsive to terrorism situations. Furthermore, to respondents in Segment 1 the attempts at reducing anxiety do the opposite—they increase anxiety.

8. Segment 2 is strongly responsive to the different terrorism events as anxiety increasers (Table **4**). The correlations greater than 0.30 are shaded; these are the combinations of terrorism incident and remedy where the remedy actually decreases anxiety in more than 30% of the cases.

Table 4: Correlation between Different Types of Terrorism Actions (Columns) and Anxiety Reduction by Remedies.

	A dirty nuclear bomb set off…	**A bomb under your car…**	**Bombs blowing up in the middle of a building…**	**Contamination of the food supply…**	**A deadly disease like smallpox or Anthrax let loose.**	**A bomb threat for a building that is a false alarm…**	**Fire raging through a building…**	**The media talking about potential Terrorism acts…**	**A Computer virus let loose…**
You believe that the Centers for Disease Control will keep you safe	0.41	0.39	0.33	0.27	0.33	0.22	0.20	0.18	0.12
You think that your local hospital will keep you safe	0.41	0.39	0.33	0.31	0.31	0.20	0.20	0.18	0.08
You believe that Homeland Defense will keep you safe	0.41	0.37	0.37	0.29	0.29	0.20	0.18	0.16	0.10
The media will keep you informed	0.41	0.37	0.33	0.29	0.24	0.27	0.18	0.14	0.12
You trust that God will keep you safe	0.39	0.33	0.33	0.24	0.29	0.20	0.20	0.20	0.08
You need to contact your friends and family to make sure they are OK…	0.41	0.35	0.31	0.31	0.29	0.14	0.12	0.16	0.12
You think that your local police will keep you safe	0.37	0.33	0.31	0.27	0.20	0.16	0.14	0.14	0.08
You believe that international cooperation in the United Nations will keep you safe	0.37	0.33	0.29	0.24	0.20	0.16	0.12	0.14	0.14
You think United Nations Forces will keep you safe	0.27	0.29	0.20	0.20	0.20	0.12	0.10	0.12	0.04
Serious terrorism event ⇐====================================⇒ Not serious event									

Note. Correlations above .30 show specific terrorism events whose ensuing anxiety can be ameliorated. The correlation was run only with the respondents in Segment 2, who showed anxiety resulting from specific terrorism acts.

9. Some terrorism incidents, such as a "dirty bomb" or a "car bomb," can be addressed by government actions. Not all remedies work, but a number do. For these situations, either the terrorism incident is tractable, or perhaps so distant in the respondent's mind that there is no problem quelling the anxiety. Thus, just because a terrorism incident is perceived to cause a lot of anxiety (*e.g.,* bombing) does not mean that this anxiety is intractable. Data in Table **3** show that the bombing causes the greatest anxiety, while data in Table **4** demonstrate, in turn, that the anxiety caused by the bombing can be reduced by many remedies.

10. However, some reasonably serious terrorism events, such as a contaminated food supply, generate anxiety that can be only reduced by a limited number of government activities, such as a better local hospital.

DISCUSSION AND CONCLUSIONS

Although public opinion research has a long and venerable history, the psychophysical "way of thinking" makes a new contribution to the field. Psychophysics looks for relations between variables, not only relations that are established by statistical analyses but also relations that are engineered by experimental design. In a sense, by importing and modifying psychophysics to public opinion research one may go from a descriptive science to an experimental science.

Such direct thinking about relations between variables is a hallmark of today's "modern psychophysics". The psychophysical methods allow the respondent to act as a measuring instrument. Psychophysics enters with its worldview and tools when there is an objective physical continuum against which these responses can be regressed, to develop a quantitative relation or "model". The key advance in this chapter is that the independent variables are not necessarily related to each other, but rather represent qualitatively different alternatives, so the relation is not between two variables (*e.g.,* sweetness *versus* sugar concentration), but rather between one dependent variable (*e.g.,* level of anxiety) and the presence/absence of the different qualitative variables (*e.g.,* different statements or messages about terrorism acts, feelings, situations and attempts at anxiety reduction). Despite the change in the nature of the model, from a continuous model to a discrete model, the psychophysical way of thinking still applies.

Public opinion and consumer researchers are accustomed to relatively large-sized samples with which to work, although the use of focus groups for political research has been gaining acceptance (Calder, 1977; Krueger and Casey, 2000). Rarely, however, do researchers talk about the very small samples of respondents, such as base sizes of one or two. The history of public opinion and consumer research focuses on the so-called *nomothetic* rules, rules that apply to large numbers of people, rather than on the *idiographic* rules, rules that apply to one person.

Ethnography and clinical psychology deal with small numbers of people, even with as few as one person, trying by observation to weave a story that applies to that one person, but at the same time has the potential to apply to many. Such small samples are perfectly acceptable in these two fields and in most observational research, simply because these observational methods do not purport to have quantitative results.

The research approach presented here lies in the middle of the world of nomothetic *versus* idiographic. The base size can be down to an "N of 1". However, the approach is not observational but rather quantitative. The experimental design applies to that N of 1 and the rules are every bit as quantitative as if the base size were 100 or 1,000 or 1,000,000 or more. The coefficients in the model represent the numerical impact of the specific phrase as a driver of anxiety for that one person. Adding more people is not to obtain a percentage, but rather to refine that numerical estimate of the impact of the element.

PROSPECTS AND OPPORTUNITIES

We'd like to end this chapter with a prospect that we feel to be exciting—the prospect of creating a public policy "data and actions" shelf of knowledge.

Based upon the approaches presented here, we see that psychophysical thinking changes the way we think about social issues, moving us from looking at patterns to looking in a more engineering-oriented way to relations between variables. The tools for social research are already available. The use of experimental design, Internet-based research and automatic analyses make the electronic bookshelf of data already feasible. The execution of the whole program is reasonable, feasible and has already been done in part. So to answer the question—we end up looking at new "Peaks in Darien," new worlds of knowledge about society, the market and about the consumer, the citizen and the person. More importantly, we end up with the prospect of new technology-enabled sciences about each of the foregoing.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of and permission by IT!® Ventures, LLC to access the data from the Deal With IT!® database and would like especially to thank Hollis Ashman for her ongoing insights.

REFERENCES

Albrecht, T. L. (1996). Defining social marketing: 25 years later. *Social Marketing Quarterly, 3,* 21–23.

Annas, G. (2002). Bioterrorism, public health and civil liberties. *The New England Journal of Medicine, 346*(17), 1337–1342.

Box, G. E. P., Hunter, J., & Hunter, L. S. (1978). *Statistics for experimenters.* New York: Wiley and Sons, Inc.

Calder, B. J. (1977). Focus groups and the nature of qualitative research. *Journal of Marketing Research, 14,* 353–364.

Donovan, R. J. & Henley, N. (1997). Negative outcomes, threats and threat appeals: Widening the conceptual framework for the study of fear and other emotions in social marketing communications. *Social Marketing Quarterly, 4*(1), 56–67.

Eisenman, D. P., Glik, D., Ong, M., Zhou, Q., Tseng, C.-H., Long, A., Fielding, J., & Asch, S. (2009). Terrorism-related fear and avoidance behavior in a multiethnic urban population. *American Journal of Public Health, 99,* 168–174.

Gofman, A. & Moskowitz, H. (2010). Isomorphic Permuted Experimental Designs And Their Application In Conjoint Analysis. *Journal of Sensory Studies,* 25(1): 127-145.

Krueger, R. A. & Casey, M. A. (2000). *Focus groups: A practical guide for applied research.* Thousand Oaks, CA: Sage Press, Inc.

Lum, C., Kennedy, L. W., & Sherley A. J. (2006). Are counter-terrorism strategies effective? The results of the Campbell systematic review on counter-terrorism evaluation research. *Journal of Experimental Criminology, 2*(4), 489–516.

Marshall, R. J., Petrone, L., Takach, M., Sansonetti, S., Wah-Fitta, M., Bagnall-Degos, A., & Novais, A. (2007). Make a kit, make a plan, stay informed: Using social marketing to change the population's emergency preparedness behavior. *Social Marketing Quarterly, 13*(4), 47–64.

Moskowitz, H. R., Porretta, S., & Silcher, M. (2005). *Concept research in food product design and development*. Ames, IA: Blackwell Professional.

Moskowitz, H. R. & Gofman, A. (2007). *Selling blue elephants: How to make great products that people want before they even know they want them.* Upper Saddle River, NJ: Wharton School Publishing.

Moskowitz, H., Gofman, A., Tungaturthy, P., Manchaiah, M., & Cohen, D. (2000). Research, politics and the web can mix. Considerations, experiences, trials, tribulations in adapting conjoint measurement to optimizing a political platform as if a consumer product. In R. Brookes (Ed.), *Marketing research in a.com environment* (pp. 223–243). Amsterdam: ESOMAR.

Rosenthal, R. (1976). *Judgment studies: Design, analysis and meta-analysis.* Cambridge, MA: Cambridge University Press.

Smelser, N. J. & Mitchell F. (2002). *Terrorism: Perspectives from the behavioral and social sciences.* Washington, DC: National Research Council.

Stevens, S. S. (1975). *Psychophysics: An introduction to its perceptual, neural and social prospects.* New York: Wiley and Sons, Inc.

Tanielian, T. L. & Stein, B. D. (2005). Understanding and preparing for the psychological consequences of terrorism. In D. Kamien (Ed.), *Emergency management, public health and medical preparedness—The McGraw-Hill homeland security handbook* (pp. 689–701). New York: McGraw-Hill.

PART IV

ADVERTISING RESEARCH AND BRAND COMMUNICATIONS

CHAPTER 13

Consumer-Driven Advertising Research

Jeffrey Ewald*

Optimization Group, Inc., Ann Arbor, Michigan, USA

Abstract: The development of breakthrough, impactful advertising historically has been considered one of the business world's more mystical creative acts—a belief that is protected and nurtured by many of the world's leading agencies. Yet many of these same agencies were early proponents of involving the consumer in the advertising development process. Clearly, they recognized the contribution that consumer insight could have on the ultimate effectiveness of advertising. The dynamic tension between an overarching mystical creative or scientific-based research philosophy is resolved differently at various advertising agencies. This chapter reviews the major types of research frequently used to develop consumer insights during the advertising development and evaluation process. The reader will develop an understanding of how different approaches and classes of methodologies contribute to the end communications deliverable and develop an appreciation for both the creative and research-oriented schools of thought.

Keywords: Advertising research, consumer, model of advertising effect, rule developing experimentation.

INTRODUCTION

The development of breakthrough, impactful advertising has historically been considered one of the business world's more mystical creative acts—a belief that is protected and nurtured by many of the world's leading agencies. Yet many of these same organizations were early proponents of involving the consumer in the advertising development process—clearly, they recognized (and continue to recognize) the contribution that consumer insight have on the ultimate effectiveness of advertising.

Despite the rhetoric of support for consumer research, there remains a dynamic tension between those who believe that "great advertising" is essentially an

***Address correspondence to Jeffrey Ewald:** Optimization Group, Inc., Ann Arbor, Michigan, USA; Tel: (313) 964-2841; Fax: (734) 527-6105; Email: jewald@optimizationgroup.com

inspirational creative act and those who hold a scientific or research-oriented perspective and believe that a strategic and logical process will improve the quality of the advertising product.

One of the major applications areas of marketing research is in advertising (Kinnear and Taylor, 1983). For example, 2008 advertising expenditures in the USA were approximately $142 billion (TNS Media Intelligence, 2009). Advertising research is used in a number of aspects of advertising, including the measurement of media audiences and testing the effectiveness of advertising messages.

This chapter looks at the advertising development process, at the research tools that more directly impact the "creative product" (*i.e.* messaging) and suggests that the emerging discipline of rule developing experimentation (RDE) may well provide a path to help balance the strategy *vs.* creativity debate.

ADVERTISING AND CREATIVITY

One philosophical point of view defines "the best creative messages" as the ones that sell the product or service. David Ogilvy (1983) said it well in his landmark *Ogilvy on Advertising*: "When I write an advertisement, I don't want you to tell me that you find it 'creative.' I want you to find it so interesting that you buy the product".

At the other end of the continuum, there are those who believe that "great creative work" is art. Among those who embrace this perspective, it is generally felt that the "quality" of the "creative art" is not what it used to be. Ad historian Stephen Fox (1985) writes: "The creative revolution of the 1960s … [was] replaced by a shift from creative emphasis to management science … from art, inspiration and intuition to research numbers". And former agency executive William Weilbacher (1993) is more emphatic about the dismal state of advertising creativity: "…the plain fact is that creativity in modern advertising is just not as good as it used to be".

In a study designed to identify whether the observed decline of creativity was anecdotal or had some empirical basis, Reid, King and DeLorme (1998) conducted a survey among top-level agency creatives. The overwhelming majority viewed "modern advertising" as more creative than it was when they entered the business, although it was felt that certain aspects of advertising creativity had changed. To facilitate their research, the investigators developed a working definition of *advertising creativity* (a major contribution in and of itself): "We define advertising creativity as original and imaginative thought designed to produce goal-directed and problem-solving advertisements and commercials". In discussing this definition, the authors noted that "advertising, as a special form of creativity, differs from artistic expression and other forms of creativity-for-the-sake-of-creativity in that originality and imagination must operate within a goal-directed and problem-solving context" (Reid, King and DeLorme, 1998).

In a review of the literature investigating creativity, it was noted that many contemporary researchers studying advertising creativity have landed on similar notions related to the idea that *advertising creativity* must balance originality with accomplishment of a goal and that a framework proposed in 1993 by Runco and Charles (the "Originality–Appropriateness" model) has become the most widely accepted (Koslow, Sasser and Riordan, 2003).

The Operationalizing the Originality–Appropriateness framework requires making judgments on what is deemed original and what is deemed appropriate. In several studies reviewed in Koslow, Sasser and Riordan (2003), it was found that people with at least some basic advertising knowledge or experience could agree on what is original. However, it turns out that Appropriateness is more context-dependent—it varies from person to person and is related to the specific objectives of a given advertising execution. In a study across creative and noncreative agency executives, it was found that "creatives tend to perceive advertising as more appropriate if they are artistic, but account executives tend to perceive advertisements as more appropriate if they are strategic". Furthermore, agency creatives did not exhibit an unbridled desire to be creative-for-the-sake-of-being-creative, believing that "being original within the confines of a tight strategy is perceived as the most creative" (Koslow, Sasser and Riordan, 2003).

If one accepts the Originality–Appropriateness framework, then the answer to the Strategy *vs.* Creativity conundrum is clear—both attributes (*i.e.* Original/Creative and Appropriate/Strategic) are critical to development of advertising that generates business success. This is confirmed in practice. A study conducted across 200 of the "most awarded" commercials in the world between 1994 and 1995 and presented at the 43rd International Advertising Festival at Cannes in 1996 was unequivocal:

> The evidence is overwhelming that well-focused commercials that are based on the right message and, in addition, deliver that message and translate it freshly, charmingly, engagingly and intelligently work better than commercials with the right message but which lack these creative qualities. Commercials with award-winning qualities are 2.5 times as likely to be associated with business success as are average commercials (Gunn, 1998).

CONSUMER-DRIVEN ADVERTISING? WHY SHOULD WE LISTEN TO THEM?

Peter Drucker offers the short answer to the above question: "There is only one valid definition of business purpose: to create a customer" (Krames, 2008). Or, as David Ogilvy (1983) so eloquently put it: "The consumer is not a moron, she is your wife".

Within an advertising-centric perspective, Ogilvy's admonition is important. In practice, those responsible for developing advertising attempt to balance many factors emanating from many stakeholders: *e.g.,* the client's organization, the product's realities, channel partners' needs, competitors' positioning and of course, the advertising agency's own internal philosophy on what makes "great advertising". And in the process of finding that balance point between these sometimes competing interests, the consumer's point of view can be diluted, if not lost.

The challenge, then, is to answer the "Appropriateness" question from the consumers' perspective. Said differently, what can we (the advertiser) say to the consumer that will result in an appropriate response by the consumer, consistent with the goals and objectives of the advertiser?

In most agencies, the need to answer this question is institutionalized through the development of the creative brief. This document summarizes what can be a painstaking analytical process involving many agency and client departments and, which is usually supported through many types of marketing research. In general, the creative brief covers these core areas:

1. Define the objective for the advertisement.
2. Select an appropriate effects model to organize thinking and information.
3. Identify the specific target audience/segment.
4. Develop a unique competitive positioning for the advertiser's product/service.
5. Provide key insights as to how the targeted consumer should/react to the message.
6. Identify critical information that must be communicated to achieve the objective.

This information, when tightly organized and packaged, defines the "creative sandbox"—the boundaries that define Appropriateness for a given assignment and within which the creative team will strive to develop one or more Original solutions.

Whereas most people conventionally assume that advertising's task is to increase sales or market share, this is not necessarily true. For example, advertising can be used to support premium pricing, or to respond to competitive activity. An in-depth understanding of the various ways that advertising can impact the target audience and the range of effects that are possible, is helpful in working with clients to define specific objectives for each ad or campaign.

MODELS OF ADVERTISING EFFECTS

Vakratsas and Ambler (1999) reviewed over 250 journal articles and books to establish what is known about how advertising affects the consumer—how it works. The paper deduced a taxonomy of models and synthesized five generalizations about how advertising works.

To develop their ad effects taxonomy, Vakratsas and Ambler first classified the many published models into two groups: those that specify *intermediate effects* (*i.e.* those that impact a consumer's beliefs or attitudes) and those that specify *direct effects* (*i.e.* purchase behaviors). Intermediate effects were then classified in three primary dimensions: cognition (thinking), affect (feeling/emotional) and experience, which can act as a feedback loop that modifies the initial cognition state, affect state and future behaviors (Fig. **1**).

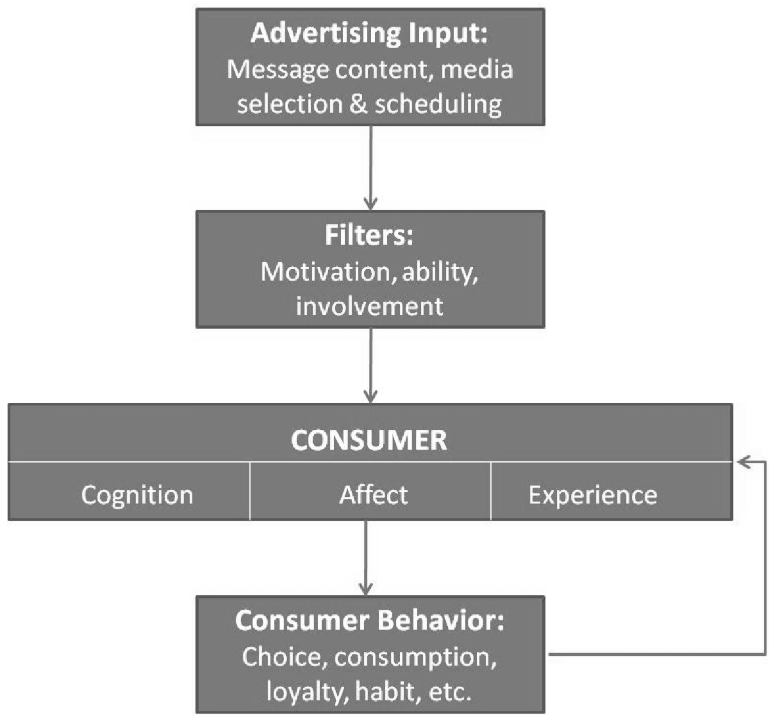


Figure 1: A framework for studying how advertising works (Vakratsas and Ambler, 1999).

Using this framework, Vakratsas and Ambler then classified the different models and theories that were reviewed. Table **1** summarizes the various models, building from models that establish direct behavior effects (assuming no intermediate effects) to those that assume only one type of intermediate effect (cognition or affect, respectively). They then describe advertising effects models that assume more than one type of intermediate effect in specific order of effects or hierarchies. To complete their classification system, the authors note that some published advertising models establish more complex hierarchies of effects depending on the context of a specific brand/category analysis ("integrative models") and some do not propose an effects hierarchy at all.

Table 1: Taxonomy of Models of How Advertising Works (Vakratsas and Ambler, 1999).

Model Type	**Sequence of Effects**
Behavioral—(market response)	"Do"—No intermediate advertising effect considered. Assumes that there is a direct behavioral effect
Cognitive information	"Think"
Pure affect	"Feel"
Persuasive hierarchy	"Think"→"Feel"→"Do"
Low-involvement hierarchy	"Think"→"Do"→"Feel"
Integrative	Hierarchy not fixed, depends on product and involvement
Hierarchy-free	No particular hierarchy of effects was proposed

A particularly noteworthy takeaway from this analysis is that it establishes the need for more balanced and holistic approaches to measuring the effectiveness of advertising:

> In summary, the evolution of models from relatively simple (Cognitive) to more complex (Cognitive, Affective, Experience) has shown the persistent significance of all three key effects and suggests that omission of any one is likely to overstate the importance of the others. Our key conclusion, therefore, is that all three effects should be included consistently in studies of advertising effectiveness … the omission of any one can lead to overestimation of the effect of the others. (Vakratsas and Ambler, 1999).

Barry's assessment of the many advertising effects models is that "advertising generally…contributes to the entire consumer behavior process—cognition, affect and conation—where the ultimate outcome is the intended behavior desired by the advertising's sponsor(s)" (Barry, 2002). These generalizations have much in common. Conation, or "impulse buying", could be considered a subset of the broader experience effect previously discussed. Both assessments also agree that intermediate effects can lead to an ultimate behavioral change.

Barry (2002) recognizes that not all advertisements have the same effect on all consumers reached and that, in fact, different effects should be planned for different audience segments. He states: "The concept of segmentation tells us that audiences are indeed different, even within relatively homogeneous segments.

Every advertising message that reaches an individual consumer effects that consumer differently, based in part on that consumer's predisposition to the product category and/or the brand".

ADVERTISING MEASUREMENT: A COMPLEX SYSTEM

Although the Vakratsas and Ambler advertising effects model is conceptually simple and intuitive, layers of complexity are introduced when one considers the myriad combinations of filters (individual motivations, product category involvement type, ability to purchase, *etc).* and consumer segments, which may occur for a given brand. "There is considerable support for a multi-path approach…namely, different people respond to different advertisements in different ways" (Vakratsas and Ambler, 1999).

This multiplication of potential effects scenarios results in a rich and complex system of effects, which needs to be measured when one is interested in analyzing every permutation of "how an ad works". It has been recognized that the inherent complexity of the advertising effects system and the lack of development in new measurement techniques appropriate for these complex scenarios have resulted in a set of effects models that lack the unequivocal validation that both practitioners and scholars desire (Barry, 2002).

Given the variety of possible effects outcomes that exist for a single ad, it is unreasonable to suppose that a single copy-testing method would be equally appropriate for all such possibilities. This condition is, perhaps, the underlying reason that a multitude of advertising testing approaches have been developed and deployed in the marketplace. To better understand how various copy-testing tools fit into the advertising development process, one should understand the typical steps followed in developing a rough idea into a finished advertising execution. Ostlund (1978) codified these steps in the following schema presented on Fig. **2**.

It is against this framework that we can understand how various advertising research methods contribute to the development and assessment of a specific execution.

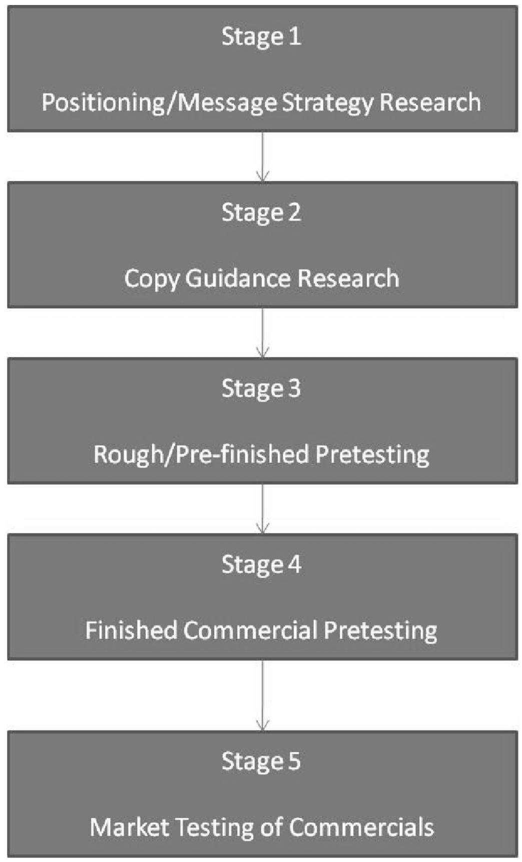


Figure 2: Copy development scheme (Ostlund, 1978).

ADVERTISING RESEARCH METHODS

A detailed treatment of every research method that has been applied to the development or evaluation of advertising messaging/copy is clearly too large a topic for just one chapter in a book; there are many excellent tomes dedicated to the subject. Interested readers might reference *The Advertising Research Handbook* (Young and King, 2008) and *How Advertising Works: The Role of Research* (Jones, 1998a) to better understand the historical perspective of advertising research and an overview of key commercially available research methodologies.

Table **2** summarizes the types of advertising research typically employed at each stage of the advertising development process.

Table 2: Advertising Research Approaches by Copy Development Stage.

Copy Development Stage	**Typical Research Approaches**
Positioning and message strategy Copy guidance research	Attitude and usage surveys Perceptual mapping Target audience research Qualitative research Competitive analysis
Rough/pre-finished pre-testing Finished ad pre-testing	Diagnostic (attitude statements) Biometric feedback Recall Persuasion (pre–post change in intent) Liking
In-market testing	Sales tracking (single-source data, *e.g.,* IRI, Nielsen) Statistical effects modeling Attitudinal tracking studies (surveys)

Research supporting the positioning strategy and copy guidance stages focuses on answering the question "What should we say?" Various forms of qualitative research often generate initial thinking, explore concepts and identify attributes that might be relevant to the development of alternative positioning statements.

The most common methods used in qualitative advertising research are projective techniques, one-on-one interviews and focus groups (Slater, 1998). Other qualitative techniques that have become popular in recent years include ethnography and means–ends research (laddering). Ethnography uses detailed "field observations" of target consumer groups to develop a "thick description" of the lived consumer experience and help address the inherent problem that people do not always do what they say (Elliott and Jankel-Elliott, 2003). The means–ends approach is an umbrella term that refers to a set of methods for interviewing consumers about the reasons for their decision choice and interpreting their responses in terms of linkages between outcomes (Olson and Reynolds, 2001). This technique is particularly useful in developing hypotheses about how attitude formations (cognitive, emotional and experiential) tie to various behavioral outcomes (see Fig. **1**).

The "positioning era" can be traced to an article on the subject published by Jack Trout (1969). In 1972, Al Ries and Jack Trout published a series of articles on the topic in *Advertising Age.* But it was their 1981 bestselling book, *Positioning: The Battle for Your Mind* (Ries and Trout, 1981)*,* that firmly established and popularized the concept on Madison Avenue (Ewald, 2009). Today, the concept of positioning is embraced by the contemporary marketing mainstream.

The term "positioning research" describes any number of techniques by which marketers try to create an image or identity for a product, brand, or company in the mind of a target audience. Development tools for positioning include many of the qualitative approaches previously discussed. In addition, conducting a thorough analysis of competitive perceptions is critical, because the concept of positioning is that what matters most is how potential buyers *perceive* the product/services as expressed *relative* to the position of competition. Popular tools to assess positioning include graphical perceptual mapping, market surveys and certain statistical techniques (http://www.valuebasedmanagement.net). The perceptual map is an expository graphical device that presents brands in juxtaposition to their competitors, according to defined criteria (Jones, 1998b). These criteria primarily relate to the way consumers perceive the brands. Understanding these perceptions enables the advertiser to discover the uncommon (*i.e.* unique) and salient qualities that should be embodied in the advertising proposition.

Research supporting the rough and finished advertising pre-testing stages generally answers the questions "How should we say or show it?" and "Are we saying (showing) it correctly?" It is interesting to note that the tools commonly employed to evaluate rough executions are the same as those used to evaluate finished executions. This has not always been the case; experiments conducted to compare key "report card measures" across the same message executed at different levels of "finish" have amply documented that rough advertising prototypes generally predict the reaction to their finished counterparts on the basic or key evaluative measures (Pierce, 1998).

Research designed to evaluate executions is commonly called "copy-testing"; there are many well-known syndicated services as well as custom types available in the marketplace. Much has been written about copy-testing, particularly the syndicated approaches. In addition, a number of large-scale validity studies have concluded that copy-testing can, when the appropriate measures are selected, be valid as a predictor of intended effects (Hoogerbrugge, 1999a; Haley and Baldinger, 2000).

Young and King (2008) categorize commonly used copy-testing measures into two types: "report card" measures and diagnostics:

Report card measures:

1. Attention or recall
2. Motivation or persuasion
3. Liking
4. Composite measures

Diagnostic measures:

1. Open-ended questions
2. Rating statements
3. Moment-to-moment tracking
4. Biometric feedback techniques

In practice, the report card measures tend to be used to help managers select between executions, or to assist in making a "go/no-go" decision regarding a specific execution. The diagnostic measures are used to better understand what an execution may be communicating (cognitively and/or emotionally), to identify which parts of an ad are "working," and to identify opportunities to improve the communications value of a specific execution, respectively.

With so many testing techniques available and with so many potential measurements that can be taken, it is critical that "a good pre-test should carry out measurements relevant to the formulated advertising objectives. A good copy-test is based on the explicit advertising framework model used" (Hoogerbrugge, 1999a).

Finally, research during the market testing stage should answer the question "Is the advertising working?" or "What is working? And what is not?" The advertising tracking study is an example of research that is frequently conducted during the test marketing phase for a new product or campaign. In addition, tracking studies have become commonplace as a measurement tool to continuously track consumer attitudes, brand perceptions and sometimes self-reported behaviors. In this advertising context, advertising tracking study questionnaires tend to follow a conventional pattern. Measures usually include some or all of the following (Feldwick, 1998):

1. Brand questions:
   a. Spontaneous and prompted brand awareness
   b. Claimed purchase behavior
   c. Brand attitude or brand image scales
2. Advertising questions:
   a. Spontaneous and prompted recall of having seen advertising for a brand
   b. Recognition of an advertisement from a visual prompt
   c. Recall of specific advertising content/message
   d. Attitudes to the actual advertising execution

Surveys constructed along these lines are typically fielded on a continuous basis or on a "pre–post" basis. The analysis can compare the continuously fielded advertising inputs (*e.g.,* creative executions, media weight, media selection) with the advertising outputs (*i.e.* the tracking measures). The analysis techniques can be as simple as graphic inspection of the data or statistically based or econometric procedures such as correlation analysis or regression. The results show the effect of the advertising. "Pre–post" studies are usually conducted as waves, with the first wave conducted before a campaign is introduced and the post-wave after the campaign has concluded, or at a predetermined point in time. Differences in measures that occur between the two waves are attributed to the advertising that "ran" during that time frame.

Advertising tracking studies can function as barometers, generating data that can help diagnose the effectiveness of specific elements of the message, media inputs, *etc.* In addition, as the time series develops, the data help establish guidelines for when new campaigns need to be developed, how competitive activity is impacting the brand, *etc.* (Hoogerbrugge, 1999b). However, tracking studies are not useful in predicting which advertising executions will generate the intended effects. Tracking studies can only tell you what has already happened. And so, ideally, tracking studies should be part of an integrated research system in which the different research approaches interconnect and mutually reinforce the specific goals established during the planning phase and articulated with a clear statement of intended adverting effect (*i.e.* desired response).

WHAT CONSUMER INSIGHTS DO CREATIVES SEEK?

As noted earlier, agency creatives do not exhibit an unbridled desire to be creative-for-the-sake-of-being-creative. Rather, they believe that "being original within the confines of a tight strategy" is when they are at their most creative (Koslow, Sasser and Riordan, 2003). In subsequent quantitative research, the authors measured how component parts of strategy, artistry and perceived originality impact the subjective evaluation of creativity by different job functions within advertising agencies. A somewhat unexpected finding was that the creatives within advertising agencies place a great deal of importance on the strategy: "The fact that strategy is very important, along with originality and artistry, demonstrates the notion that creatives require strategy as stimuli in the form of a brief, as much as they need artistry and originality" (Koslow, Sasser and Riordan, 2003). (As a point of contrast, it was found that account executives focus on strategy and will settle for unoriginal but artistic "craft" in advertising in order to please the client).

There are two types of insights typically provided to agency creatives by researchers and brand planners (Maloney, 1998):

1. Brand positioning statement: brand positioning can be defined as the way the customer should think about the brand relative to competitors.

2. Ad strategy: provide guidance and direction for the development of the brand's advertising campaign. The ad strategy comprises the Who, What and Why in addressing a specific brand issue or objective.

These insight documents can be presented separately or together as part of an overall creative brief that, in a disciplined agency, occurs before significant investment of the creative department's time and other resources. So, to some extent, it appears that agencies are typically delivering the strategy definition (*i.e.* the brief) that creatives say that they want.

However, the results of a study presented at the 2003 ESOMAR Congress (Moskowitz *et al.* 2003) suggest that creatives actually want more than strategic consumer insight—they also desire hard data to test their hypotheses and drive actionable decisions. This finding was developed from an ingoing assumption that "people don't know what they want, but they will know it when they see it". This assumption is often the *raison d'etre* for conducting a conjoint study (Wittink and Cattin, 1989). One of the beauties of conjoint measurement is that it is relatively impervious to "politically correct" answers. With the elements varying in mix/match combinations, it is hard to identify the appropriate answer. Consequently, most answers are intuitive, rather than carefully reasoned.

Table **3** reveals that creative people (in this case, graphics designers) respond positively to promises of consumer insights that are (1) predictive, (2) quantitative and (3) drive actionable decisions. (They also respond to being able to better understand how technology impacts customers' decision making). Because creatives are responsible for producing a tangible piece of communication, they clearly want hard data from which to make decisions about what works and what does not.

Table 3: Select Strong Message Elements: Conjoint Responses by Job Type.

Among Creatives (Graphics Designers)	Among Strategic Brand Planners
After all, actionable consumer insight isn't about what consumers think TODAY…it's about what they will think TOMORROW	What motivates your customers to make a choice? Cost, convenience, habit, or something else?
Anyone can provide qualitative insights…we deliver hard data to test your hypotheses and drive actionable decisions	Consumer insight is a blend of art and science…we strive to provide the best of both
A partner that can help you understand how technology is changing the way your customer makes decisions	Learn what products your customers use…the brands they prefer…

In contrast, brand planners respond to promises of obtaining a better understanding of consumer choice—whether it is a choice of product or brand. This too is directly in line with their job function, as planners are responsible for developing strategic insights and preparing the creative brief.

This conjoint study is consistent with previous research suggesting that agency creatives value strategy and the tight definition of task and desired consumer effect that is imposed by the creative brief. However, it differs from previous research that suggested that creatives are less favorable toward (and even feel threatened by) having their creative output evaluated quantitatively (*i.e.* copy-testing; Morgan, 1984/1985; Vaughn, 1982/1983). The discrepancy can be explained by looking at where the quantitative measurement occurs in the advertising development process. Whereas Morgan and Vaughn found that creatives do not like their OUTPUT evaluated quantitatively, Moskowitz *et al.* (2003) discovered that creatives evaluate quantitative data that help test hypotheses and drive actionable decisions—that is, data that can help creatives determine direction while still in the developmental phase.

Fortunately, a quantitative approach that bridges between traditional positioning development research and rough copy development research has emerged on the scene. In the next section, we will introduce RDE and its role in helping define the messages that best resonate with the target audience(s) to achieve the desired end-goal for a specific commercial.

AD RESEARCH USING RULE DEVELOPING EXPERIMENTATION

While not a totally new approach, RDE has been best codified by Moskowitz and Gofman in their best-selling book *Selling Blue Elephants* (Moskowitz and Gofman, 2007): "RDE is a systematized solution-oriented business process of experimentation that designs, tests and modifies alternative ideas…in a disciplined way so that the developer and marketer discover what appeals to the customer, *even if the customer can't articulate the need, much less the solution*!"

As discussed earlier, advertising creatives are less interested in testing their final executions (the dreaded "beauty contest"), but welcome and value anything that helps them focus their creativity on solving the communications challenge. RDE applied to messaging defines "the sandbox". Applied at the early stages in the creative development process (Stages 1 and 2), RDE identifies *what to say, how to say it and to whom to say it*.

As applied to the development of messaging, the RDE process follows these straightforward steps (adapted, with permission, from Moskowitz and Gofman, 2007):

1. Think about the communications challenge and identify a wide range of "things we could say or show" about the product or service of interest. These "things we could say or show" will become the core inputs into the RDE process. In many cases, the "things we could say or show" will emerge from background research and competitive analysis often done by the account planner and might follow proven communications architectures such as brand name, product/service features, product/service claims, rational consumer benefits, emotional consumer benefits, price or offer, call to action, *etc.* (This is the most difficult part of the RDE process and is where expertise and familiarity in the product category and in developing marketing communications comes in).

2. Mix and match the elements according to a special experimental design to create a set of "rough ad" prototypes. (This step is usually done automatically by a tool).

3. Show the prototype "rough ads" to consumers. (This step can be done *via* an Internet survey or in centrally located focus group facilities).

4. Analyze the results using a regression module. The magic of experimental design estimates the contribution of each of the individual message stimuli (inputs) to the desired consumer response (output). The desired consumer response is defined by the team in advance and is based on the specific goal of each advertisement. Examples of typical desired responses: increased intent-to-buy, improved price–value perception, likelihood of visiting my local dealer, *etc.*

5. Optimize the "rough ad" or communications platform. To uncover the optimal message, you just need to find the optimal combination of elements that generate the highest score.

6. Identify naturally occurring attitudinal segments of the population. These segments span traditional demographic group and cluster respondents on the basis of how they respond to different messages. (For example, for some product categories, we have found that people who shop over the Internet respond to different messages compared with those shoppers who prefer bricks-and-mortar retailers. Knowing this helps to increase the desired response by 10–50% or more).

7. Apply the generated rules to the development of creative executions. The agency team now has a tool—specific to their client's product/service—that allows them to "dial in" various parameters and to immediately generate the messaging platform, which will best achieve the desired goal among target audiences.

Implementing the RDE process into a "creative organization" can be challenging. It should be recognized that there will be some stakeholders who attempt to block (consciously or unconsciously) the introduction of yet another research tool. Commonly cited reasons include incremental cost, increases in development timelines and the fear that research will stymie the creative process and therefore the end creative product.

As is true when introducing any new process to an organization, involving those who will be impacted is an important requisite to the achievement of group acceptance and success (Mills, 2008). Having successfully introduced RDE processes to many advertising agencies (including small boutique agencies as well as large multinationals), the author experience suggests that it is critical to fully brief the entire client and agency team (including researchers, planners, account management, media planners and creatives) on the RDE process and to encourage contribution of test elements from both client and agency teams. When the client and agency jointly "own the inputs," the impact of applying RDE to the advertising development process can be profound. Creatives who have experienced and embraced RDE call it "liberating". In their view, the study results speak for themselves—the strategic "sandbox" is no longer up for debate. This frees them to apply intuition and art to the creation of copy and visuals, which best communicate the proven messaging platform in ways that will grab the viewer's attention and break through the clutter.

SUMMARY

This chapter first explored creativity in the context of commercial art such as advertising (as opposed to the fine arts). We developed the Originality–Appropriateness framework of creativity as a way of defining advertising creativity and then established the notion that specific goals against a defined target audience could operationalize the Appropriateness dimension. Next, we reviewed advertising measurement and summarized types of advertising research methodologies that are in common use. Finally, we introduced new insights as to what agency creatives want (consciously and subconsciously) from their research and planning colleagues and then introduced RDE as a tool with the potential to bridge the gap between the research support that creatives want *versus* what they currently receive.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGMENTS

None declared.

REFERENCES

Barry, T. E. (2002). In defense of the hierarchy of effects: A rejoinder to Weilbacher. *Journal of Advertising Research, 42*(3), 44–47.

Elliott, R. & Jankel-Elliott, N. (2003). Using ethnography in strategic consumer research. *Qualitative Market Research: An International Journal, 6*(4), 215–223.

Ewald, J. (2009). Getting the positioning right: Advertising planning. In H. Moskowitz, I. Saguy, & T. Straus (Eds)., *An integrated approach to new food product development* (pp. 387–399). Boca Raton, FL: CRC Press.

Feldwick, P. (1998). Tracking studies. In J. P. Jones (Ed)., *How advertising works: The role of research* (pp. 234–243). Thousand Oaks, CA: SAGE Publications, Inc.

Fox, S. (1985). *The mirror makers: A history of american advertising and its creators.* New York: Vintage Books Edition.

Gunn, D. (1998). Do award-winning commericals sell? In J. P. Jones (Ed)., *How advertising works: The role of research* (p. 276). Thousand Oaks, CA: SAGE Publications, Inc.

Haley, R. I. & Baldinger, A. L. (2000). The ARF Copy Research Validity Project. *Journal of Advertising Research, 40*, 114–135.

Hoogerbrugge, M. (1999a). Pre-testing. In G. Franzen (Ed)., *Brands & advertising: How advertising effectiveness influences brand equity* (pp. 246–247). Oxfordshire: Admap Publications.

Hoogerbrugge, M. (1999b). Tracking research. In G. Franzen (Ed)., *Brands & advertising: How advertising effectiveness influences brand equity* (pp. 250–260). Oxfordshire: Admap Publications.

Jones, J. P. (1998a). *How advertising works: The role of research.* Thousand Oaks, CA: SAGE Publications Ltd.

Jones, J. P. (1998b). Perceptual mapping. In J. P. Jones (Ed)., *How advertising works: The role of research* (p. 136). Thousand Oaks, CA: SAGE Publications, Inc.

Kinnear, T. C. & Taylor, J. R. (1983). *Marketing research: An applied approach.* New York: McGraw-Hill.

Koslow, S., Sasser, S. L., & Riordan, E. A. (2003). What is creative to whom and why? Perceptions in advertising agencies. *Journal of Advertising Research,43*, 96–110.

Krames, J. A. (2008). *Inside Drucker's brain.* New York: Penguin Group.

Maloney, C. A. (1998). *How to be a better client advertiser.* Retrieved 09.12.2009, from www.adcollege.com/dispatches/dispatch_1.pdf

Mills, C. (2008, July 8). *Leading your organization through change: A management plan—5 rules for success.* Retrieved December 13, 2009, from http://www.tagonline.org

Morgan, A. I. (1984/1985). Point of view: Who's killing the great advertising campaigns of America? *Journal of Advertising Research, 24*(6), 33–35.

Moskowitz, H. R. & Gofman, A. (2007). *Selling Blue Elephants. How to Make Great Products that People Want Before They Even Know They Want Them.* Upper Saddle River, NJ: Wharton School Publishing.

Moskowitz, H., Katz, R., Silcher, M., Flores, L., Ewald, J., & Keeling, C. (2003). A seat at the table: What business wants from providers of consumer insights. In *Proceedings of the ESOMAR Congress* (pp. 22–23,31). Prague: ESOMAR.

Ogilvy, D. (1983). *Ogilvy on advertising.* London: Pan Books.

Olson, J. C. & Reynolds, T. J. (2001). The means-end approach to understanding consumer decision making. In T. J. Reynolds & J. C. Olson (Eds)., *Understanding consumer decision making: The means-end approach to marketing and advertising strategy* (p. 3). Mahwah, NJ : Lawrence Erlbaum Associates, Inc.

Ostlund, L. E. (1978). Advertising copy testing: A review of current practices, problems and prospects. *Current Issues and Research in Advertising, 1*(1), 87–102.

Cattin P. & Wittink, D. (1982). Commercial use of conjoint analysis: A survey. *Journal of Marketing, 46*, 44–53.

Pierce, P. (1998). Rough *versus* finished commercials in research. In J. P. Jones (Ed)., *How advertising works: The role of research* (p. 170). Thousand Oaks, CA: SAGE Publications, Inc.

Reid, L. N., King, K. W., & DeLorme, D. E. (1998). Top-level agency creatives look at advertising creativity then and now. *Journal of Advertising, 27*(2), 1–16.

Ries, A. and J. Trout (1981). *Positioning: The Battle for Your Mind.* New York: Warner Books.

Slater, J. S. (1998). Qualitative research in advertising. In J. P. Jones (Ed)., *How advertising works: The role of research* (p. 127). Thousand Oaks, CA: SAGE Publications, Inc.

TNS Media Intelligence. (2009). *TNS Media Intelligence reports.* New York: TNS Media Intelligence.

Trout, J. (1969). “Positioning" is a game people play in today’s me-too market place. *Industrial Marketing*, 54(6), 51-55.

Vakratsas, D. & Ambler, T. (1999). How advertising works: What do we really know? *Journal of Marketing, 64*(1), 26–43.

Value Based Management (2008). *Creating the perception of a product/brand/company identity*. (n.d).. Retrieved December 2008, from http://www.valuebasedmanagement.net

Vaughn, R. L. (1982/1983). Point of view: Creatives *versus* researchers: Must they be adversaries? *Journal of Advertising Research, 23*(6), 45–48.

Weilbacher, W. M. (1993). *Brand marketing.* Lincolnwood, IL: NTC Business Books.

Wittink, D. & Cattin, P. (1989). Commercial use of conjoint analysis: An update. *Journal of Marketing, 53,* 91–96.

Young, C. E. & King, P. D. (2008). *The advertising research handbook,* (2nd ed).. Seattle, WA: Ad Essentials, LLC.

CHAPTER 14

The ROI of Woo: Starting, Sustaining and Improving the Relationships Business and Brands Have with Consumers

Mike Gadd* **and Jennifer Gadd**

Gadd Research Inc., Toronto, Ontario, Canada

Abstract: Our analysis of over 600 ads, campaigns, and ideas for ads for packaged goods, automotive, retail, telecommunications, financial services, and not-for-profits indicates that of all the variables and relationships examined the emotional connection with advertising—that is, how it makes consumers feel or how they want to feel—is the number one driver of purchase interest. This chapter examines how this finding might be expanded to measuring and understanding the importance of relationships consumers have with brands.

Keywords: Advertising, brand-consumer relationship, rule developing experi-menttation, CRM.

INTRODUCTION

The prologue of the book *Selling Blue Elephants*, which is about the evolution and the various applications of structured experimentation with consumers, suggests that "in order to survive, businesses must understand the current customers' needs, both current and not yet thought of" (Moskowitz and Gofman, 2007).

In order to thrive [and improve return on investment (ROI)], businesses must understand the relationships they have with current customers and targeted consumers and sustain them or improve them. In order to do this, stakeholders have to woo. Businesses should converge from marketing-centric to consumer-centric strategies (especially in deciphering the emotions and attitudes of current customers and targeted consumers), possibly to primarily customer relationship management (CRM)-centric.

According to Seth Godin, businesses can no longer be interruptive in their

***Address correspondence to Mike Gadd:** Gadd Research Inc., Toronto, Ontario, Canada; Tel: 416-362-9116; Fax: 416-362-0636; Email: mgadd@gaddresearch.com

Alex Gofman and Howard R. Moskowitz (Eds)

marketing (Godin, 1999). Our own study on research, partnered with an ad agency and media company, agrees that much advertising continues to be too interruptive and not engaging enough. It needs to be more aligned with the emotions and attitudes of the targeted consumer.

Other leaders of the industry support this direction. In the words of Jim Stengel, former World CMO of Procter & Gamble, the purpose of advertising is to start a relationship (Precourt, 2007). Kevin Roberts, CEO Worldwide of Saatchi & Saatchi, one of P&G's ad agencies, went a step further in his book *Lovemarks* by suggesting that all stakeholders need is love to create unconditional loyalty among consumers (Roberts, 2004).

THE IMPORTANCE OF THE WOO

Contemplating the meaning of the word "need" immediately brings to mind the title of the Beatles song "I Need You". The notion of getting more from a relationship, and by expansion improved ROI, is perhaps underscored in the words of another Beatles song, "The End": "and, in the end the love you take is equal to the love you make".

A relevant definition of the word "need" is "motivation: the psychological feature that arouses an organism to action toward a desired goal; the reason for the action; that which gives purpose and direction to behavior" (Wordnet Search, 2009). Looking to psychology and Maslow's hierarchy of needs (Hagerty, 1999), love is positioned firmly in the middle above safety and physiological needs but below esteem and self-actualization needs.

The human brain collects information and processes in the following pattern: sense, feel, maybe think, and do (Hill, 2005). The notion of sensing goes back to the well-developed sense of smell at the beginning of the human experience. The capability to feel, good or bad, gives support to the importance of the lateral thinking or emotions in decision making. Of course, a few companies such as Nike believe in and advocate just doing it.

Our analysis of over 600 ads, campaigns, and ideas for ads for packaged goods, automotive, retail, telecommunications, financial services, and not-for-profits indicates that, of all the variables and relationships examined, the emotional connection with advertising—that is, how it makes consumers feel or how they want to feel—is the number one driver of purchase interest and the only driver behind both the key measurements of purchase interest (a rational measure) and likeability (an emotional measure) (see Fig. **1**).

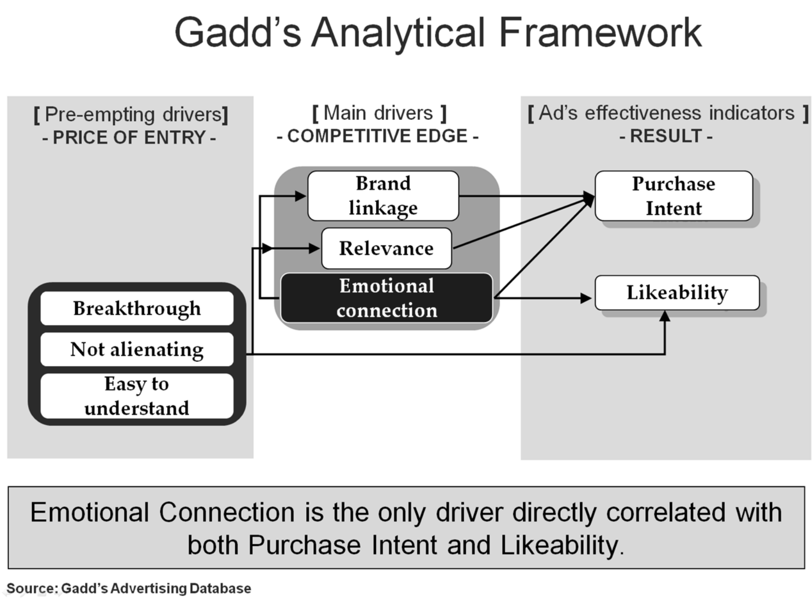


Figure 1: Drivers of advertising purchase interest and likeability.

Emotionally connecting advertising is noticeable (consumers sense it and feel good or not so good about it and screen it in or filter it out), clear (they get it), fresh (it breaks through the clutter), and makes the target consider a relationship with the advertised brand of product or service.

In their daily world, consumers are bombarded with competition for their attention or consumption. As an average grocery store might offer 25,000 SKUs (shop keeping units or products) (Katsenelson, 2004), consumers have to make choices. Shoppers may make choices quickly, or they may adopt a considered approach. Unconditional love though makes it easy to reduce the choice in-store and helps with decision making. It frees up time for the things that are important, the things on which consumers want to spend time.

BRANDS WORKING THE WOO

The results from years of market research studies reveal that there can be a large gap between brand awareness and brand purchase consideration. The brands that come to mind first are not always the most considered. Brands that are used are not necessarily the ones best loved; sometimes consumers feel locked in, perhaps abused, and express that they lack choice. Sometimes, they question the quality and value of the product. Relatively speaking, their relationships are dysfunctional and need fixing.

The American Dream of Oprah Winfrey floats on a sea of love. Coca-Cola's success is usually attributed to the connection or inspiration of target consumers with its happiness strategy. McDonald's is just "lovin' it". Wal-Mart is in the process of dialing up its reputation for caring about its customers and consumers by greening its locations and vendors to lower costs and lower everyday prices. Apple knows how to make and design products and services, such as the iPod, that make it easy and beautiful for the consumers.

Some would argue that the iPod was not the best product and not founded on the best MP3 technology. Yet iPod stands heads and shoulders above its primary competitors and the iPod success story is well documented. Why has the iPod been so successful? On the one hand, it is an evolution of the MP3 player and the revolution of a unique bundle of technological, distribution, and design benefits. On the other hand, Steve Jobs knew a thing or two about making sure that Apple products and services are simple to use, convenient to store, and cool to look at and feel. It's the design. It's the touch. In doing so, he created a cultural icon that perhaps has no bounds.

The healthiest brands simply have the healthiest relationships with their targeted consumers. They are based on unconditional love, until a better suitor such as the iPod, Google, Facebook, or Craigslist comes along to make it easier and more beautiful for them. And even then, as with the case of Coca-Cola *versus* Pepsi-Cola, love can beat all even if the product does not perform as well in blind-taste tests.

All businesses should follow suit by developing and designing products and services that are easy and beautiful to use. In their book *Selling Blue Elephants*, Howard

Moskowitz and Alex Gofman make RDE, a complex conjoint analysis research and analysis tool, easy and beautiful simply by sharing case histories and the story of Alison, whose business at every chapter thrives because of its use of RDE.

Even financial services companies might make it easy and beautiful for consumers during difficult economic times. According to our database of advertising, levels of engagement with the advertising for banks and brokerage companies are lower overall than they are for other categories.

It could be that engagement levels are lower for a variety of reasons. The challenge has been to differentiate brands of financial services. Finances should be of high interest to consumers but are not always. For many, the household accounts and discussions on finances can be more of a chore than a delight. Again, the higher scoring financial ads in our database overall plays more to the emotions and attitudes of consumers, and make it easy for them, as opposed to the features of a particular product or service, which can sound difficult.

Canadian bank TD has made it easy and beautiful for consumers simply by having a reputation for superior customer service, *e.g.,* extended branch opening hours. TD's edge comes through loud and clear in all kinds of surveys. The bank has been awarded the JD Power Retail Banking Satisfaction prize for four years in a row (TD Bank advertising and media room, 2009). TD also has announced a greenhouse emissions plan.

Another bank, local to British Columbia with headquarters in Vancouver (the location for the 2010 Winter Olympic Games), has won accolades for its involvement with the local community and its leadership with respects to the environment. These examples show that it is possible to develop initiatives to create brand love, even for the toughest of categories.

APPROACHES TO MEASURING THE WOO

The importance of business or brand relationships, as well as the need to measure them, improve them, and optimize them using appropriate research techniques and metrics is central to calculating the health of brand relationships and the ROI of marketing initiatives, *i.e.* "the love it takes". The question is what are or where are the best measurement techniques and metrics to measure the relationships brands foster with consumers?

Depending on orientation, training, philosophy, commitment, investment, and perhaps their own personal DNA or NLP (neuro-lingual programming), there are different schools of researchers.

A large school still uses self-reporting surveys and qualitative research essentially to validate more than to discover. This form of research often relies on metrics and industry standards in order to interpret data, because results on their own without contextualization can be difficult to interpret. Much of the value of this form of research comes from the associated statistical analysis that reveals hidden motivations and thoughts. Whereas this form of research certainly has its uses and has long contributed to corporate decision making, many would argue research of this type only hovers at the surface and does not go deeper into the true elements of decision making and emotions.

There is another school of practitioners who believes that deciphering the emotions and attitudes of consumers starts first by understanding consumer behavior, including transactional analysis, and working back in order to understand motivations. Their thinking is influenced by beliefs that people do not know what they want and that there is a difference between what consumers say they want and what they actually want. Some proponents of this approach believe consumers are unreliable and unable to identify their subconscious thoughts and motivations. This attitude toward self-reporting methodologies may explain the uptake of research that relies on observation rather than claimed usage. This includes methodologies such as ethnography, tracking services online, in-store and "in-person" (using eye tracking, brain scanning, biometrics, facial decoding), and the study of signs and symbols as elements of language using semiotics. Seeing, or tracking, is considered believing for proponents of this style of research.

Yet another school believes that consumers can help marketers; they just need a little help in enabling and articulating their emotions. This can be achieved through indirect research exercises, such as RDE or projective techniques that are founded in clinical psychology and sometimes used by qualitative researchers. The metaphors chosen are of interest but the real insights, and the real differences in motivations by gender, age, culture, values, and lifestyles, come from the reasons why.

An emerging school likes the idea of triangulation. Here, researchers simply believe that there can be issues with all sorts of research methodologies, and as such, they want to tackle the same problem from different angles using more than one research methodology.

With all the differing theories and methodologies within the arena of marketing research and understanding brand relationships, it is not surprising that there is yet another school of thought. That is the school that shuns research and dismisses any gathered information as inaccurate and unreliable. Many of the proponents of this thinking dismiss research as a crutch that gets in the way of individual expertise and creative thinking. Research is often blamed by this school as rewarding and recognizing only the mediocre and familiar.

Pretty much any business can woo by making its product or service easy and beautiful for consumers. According to the late Phil Dusenberry in his book *One Great Insight Is Worth a Thousand Ideas*, all it takes is a RAISE: Research, Analysis, Insight, Strategy, and Execution (Dusenberry, 2005).

For researchers who have suffered the slings and arrows of outrageous misfortune some of the time, it is encouraging that Dusenberry, former chairman and chief creative officer of BBDO North America, advocates upfront in his book that stakeholders need to place their faith in research first.

Of course, the telling word is the word “faith”. Unbelievable as it may sound, innovators do not always automatically have faith in research. Their default can be that research is the idea killer. A plaintiff cry might be: “All my best ideas never see the light of day”. Innovators might typically want to skip the first two steps to insight, strategy, or execution, because they believe that people do not know what is possible conceptually or that they are not imaginative enough to have a valid point of view.

So why did Dusenberry, creative guru behind many memorable campaigns for major advertisers such as GE and Pepsi Cola, have faith in research and ordain it? “Research comes in many forms. They’re all valid if they produce valid insights”. Dusenberry’s number one criterion was coming up with the great insight in order to provide a platform for hundreds of great creative ideas to be based on. He proclaimed: “Insights do not appear in a vacuum. They appear when you start assimilating information”.

For great ideas to follow, Dusenberry wanted the great insight at the beginning of the process. However, research can be essential not only at the beginning to understand the relationship a business or brand has with the current customer or targeted consumer, but along the way to understand both the potential impact marketing initiatives might have on the relationship and the impact they actually have.

The case has started to be made for measuring the wooing of brands to start, sustain, or improve relationships with current customers or targeted consumers. The question is what is the most accurate, reliable, and comprehensive way to measure? What are the best criteria to use?

Conventional quantitative research surveys do not measure the relationships between brand and current customer or targeted customer. For the most part, they focus on recall of and attitudes toward the marketing initiative and brand as opposed to the emotional relationship with the consumer. They are mostly marketing-centric, purporting to measure the effectiveness of the marketing, or customer-centric, purporting to measure consumer behavior. They are not customer relationship-centric researching the predictably irrational emotions that drive interest.

Typical surveys are question and answer sessions with respect to awareness, usage, recall, communication, comprehension, believability, likeability, attitudes, and purchase interest. Pre- and post-exposure or test sample *versus* control comparisons might highlight the love a marketing initiative makes.

In many respects, though, conventional quantitative research methodology reflects an older Learn Think Do (Hill, 2005) model of how the brain was thought to collect information and process it. It assumes a rational being takes the time to look at and study information on a conscious level, chews it over, and then decides on what action to take, if any.

Potentially, each model of thinking may align with different segments of consumers. The point is that measurement techniques and metrics that are designed to help with decision making—especially for stakeholders such as Jim Stengel, Steven Jobs, and Kevin Roberts—have to take into consideration either the newest model or both.

Some well-established models focus on the assets of brands. The problem here is that brands such as Microsoft, Coca-Cola, McDonald's, Wal-Mart, and Ford can have and do have quite different brand assets. They represent to varying extents of familiarity and favorability. And even their respective values can vary depending on whether conventional or projective techniques are used in the measurement.

One premise that may be worth exploring is whether measuring the relationships consumers have with businesses and brands will result in more accurate, reliable, and comprehensive metrics of brand health and ROI for evaluating and optimizing marketing initiatives across media platforms. Perhaps one metric would be a distillation of the measurement techniques used. The ultimate metric might be a blend of the relationship measurement techniques with the transactional data.

The new stand-alone or blended relationship measurement techniques could be a substantial improvement over the current ROI measurement techniques and metrics. The biggest criticism of currently available ROI measurement techniques is that they are based on short-term transactional data that do not reflect long-term effect or investment.

In order to ensure systematic, robust brand relationship ROI assessments and tracking, a quantitative research approach is required using appropriate indirect and some good direct research approaches. The research method we have in mind will be standardized for all brands, with opportunities for more detailed custom measurements applied to those particular brands that require deeper dives into brand associations or emotional cues.

The model should be aligned with the observations of some of the greatest thought leaders in marketing, advertising, and design: brand relationship measurements and the ROI metric. Success can be defined then as warming or winning the hearts and minds of consumers.

The inputs for a systematic brand relationship model could consist of some, or all, of the advanced emotions used by Robert Plutchik in his model for love of optimism, awe, contempt, disappointment, aggressiveness, and remorse, which are composed of joy, trust, anticipation, fear, surprise, disgust, anger, and sadness (Wikipedia, 2009). Depending on the nature of the brand's wooing, some emotions and their composition are likely to be considerably more effective than others.

Secondary emotions that might also be included as inputs are affection, lust, or longing. Tertiary are adoration, affection, fondness, liking, attraction, caring, tenderness, compassion, sentimentality, arousal, desire, lust, passion, infatuation, or longing.

In summary, the questions to be answered by a new model that focuses with respect to the wooing by a business or brand are:

1. What relationship do current customers or targeted consumers have with our business or brand now?
2. How can the relationship be improved?
3. What will be the impact that new marketing initiatives might have on the relationship?
4. How can marketing initiatives be optimized to strengthen the relationship? How can these initiatives be made easier or/and more beautiful for customers or consumers?

The best types of brand relationship ROI measurement might result from indirect research exercises only—such as RDE, brand personification, or metaphor elicitation—or from a combination of indirect and direct research exercises.

The brand relationship ROI metric will emerge by measuring and analyzing the influences that drive passion for a brand using appropriate statistical techniques.

The quest will be to have a standard metric across all business or brands or different leagues or divisions of businesses or brands based on slightly different criteria driving relationships. The goal of businesses or brands might be to compete in the champions division for relationships or at least the premiership.

This new metric should aim to establish whether maximizing brand relationships or love through new marketing initiatives will lead to improved customer loyalty, improved conversion rates, and improved ROI.

CONSIDERATIONS FOR MEASURING THE WOO

In order to measure the woo for a given brand, it is necessary to have a relevant up-to-date benchmark of the customer or target consumer's relationship with the brand against which a test initiative can be compared. For a more established brand, especially where there may not be much scope for movement, the nuances can be measured by analyzing the driving metrics of the Plutchik-based emotions to determine whether they have become more advanced as a result of exposure to the woo.

The notion of using a test *versus* control approach for measuring the woo may sound rather elementary to some. The approach is extremely well-grounded in scientific theory but it goes against the grain of much market research thinking over the past 100 years or so. Over this period of time, many techniques have been developed and devised to test individual elements of the marketing mix, which are compared with category or industry norms. Major market research companies have made their fortunes this way and/or with black box modeling.

Paraphrasing former researcher-cum-advertising guru David Ogilvy, marketing research has become more about validation or auditing as opposed to illumination or discovery. Based on the Boston Consulting poll of client researchers and research marketing users conducted in later 2009, most would appear to agree that research is too tactical and not strategic enough.

Measuring the woo requires a more holistic approach. This means having an up-to-date benchmark read of the relationship customers or target consumers have with the brand, against which woo initiatives can be tested and compared. Successful measurements of the woo are a stronger relationship—positive differences between test and benchmark—or a more intense passion toward the brand.

The benchmark survey can be much more than a benchmark. It can be an opportunity to define and segment customers, from those with the strongest feelings about the brand, who are generally the most brand loyal, to the least, and explain why. This process in itself presents opportunities to identify segments of customers for migration into higher value segments, thus improving ROI potential. It can also add value by serving as the basis for a Tracker. The benchmark sample size has to be sufficiently large (1,000 or more) to allow for analysis, including segmentation, and subsequent comparisons of woo surveys (150 or more) with customers or targeted consumers.

The approach overall is standardized and systemized:

1. Qualitative research is conducted first to identify the brand associations, emotional cues, metaphors, and projective exercises required in order to enable survey respondents to get in touch with and articulate their feelings.

2. The quantitative benchmark survey is conducted to define the relationship(s) and the drivers. There's much scope for analysis and insight gathering. Follow-up qualitative interviewing can be conducted, usually by tele-web, or ideate *via* mini-groups.

3. Test surveys of new relationship building or sustaining initiatives are conducted among customers and test consumers whereby key woo metrics and drivers are determined and compared with relevant benchmarks. Again, there are options for post-qualitative research deeper dives and insight gathering.

While the approach is standardized and certain aspects of the survey instrument—the associations, emotional cues, metaphors, and projective exercises—much is customized to suit the brand relationship, and the needs and objectives of the brand relationship management team, hence, the need for qualitative research upfront.

The research and measurement techniques can be conducted efficiently in markets with high Internet usage penetration. In markets with low penetration, homework exercise/extended focus groups can be used instead of ethno diaries and personal instead of online interviewing. The key consideration in researching relationships and consumers' emotions is enabling consumers, generally through indirect exercises, to connect with their feelings and articulate them. Experience shows that a picture or piece of music submitted as a metaphor or simile enables "a thousand words".

The initial qualitative research exercises can be extended to members of the marketing team, other employees (which can improve hiring, training, motivation, and retention), and creative experts. This can be useful in itself as it can provide insight in terms of how congruent the various populations are and the "stretchability" of the brand. The inputs collected—plus images, constructs, and brand asset statements from our library—are used to help develop the survey instrument that will be fielded.

The qualitative research exercises can be conducted in different countries and across cultural groups for multicountry applications. Focus groups and homework exercises can be conducted instead among those cultural groups where online access is not a viable option.

Survey inputs vary from study to study. Even the most basic questions included can vary depending on frequency of category buying, the speed of product change in the category, and the degrees of loyalty toward brands. For certain fast-moving consumer goods (FMCG) categories, consumers may pretty much use one brand most often. For others, they may have more of a repertoire of brands from which they choose. For certain brands of luxury vehicle, loyalty levels can be low because users might like to experiment with different marques or makes, or because they may have become dissatisfied. One brand usage loyalty measurement does not fit all. A brand attitude measurement, *i.e.* brand feelings, is more ubiquitous, and is appropriate for segmenting customers from those with the strongest feelings to those with the weakest feelings.

Among various cultural groups, the brand associations, emotional cues, metaphors, and projective techniques for the brand under study might very well be different. For a given brand and its communications, it may be appropriate to include brand sorts (to help understand positioning), brand parties [for measuring word-of-mouth (WOM)], or brand obituaries (for reviving a brand), and RDE for various applications. Using common brand assets and/or visual cues for all brands across cultural groups is not a recommended practice.

Suffice it to say, the analysis techniques are dependent on the measurement techniques and metrics actually used. There are many opportunities for understanding the woo, its power, its cause, and possibilities for optimization. There are opportunities for modeling! The best thing about statistical modeling is that it can simplify things. The worst thing is that it can oversimplify things and not really add any understanding or insight.

Measuring the woo underscores the differences between brands and heightens the need for measuring techniques and metrics that are not generic but are highly specific for measuring differentiated brands. Brands and the woo of brands are designed to be different. Measurement techniques need to measure the differences. Any norms or benchmarks need to be customized for a given brand to measure its woo. Think of the respective brand woo of Coca Cola (happiness), BMW (joy), and Dove (self-respect). How much sense is it to average them?

A key consideration for also measuring the woo is the degree of finish of the stimulus used in the testing. For example, for decades, the research industry has indicated rough unfinished materials test as well as finished. Based on analysis of our database of pre-tested ads, the norms or the averages are the same for finished and unfinished for all factors save for one. The emotional connection is higher for finished ads than unfinished. It would seem that casting, lighting, directing, production qualities, computer effects—the craftsmanship of creating beautiful ads that are easily understood—make the difference. This gives support to the old adage that it is not so much what is said that is important but how it is said—the woo.

Our highest-scoring ads in our database are all finished ads. The highest scoring overall increased sales 31%; it was a very humorous ad that was developed on a telling insight into the human condition, cloaked in the indisputable personality of the brand. The highest-scoring animatic ad ever has an easy-to-follow, but very clever, storyline that set the creative brief for executions to follow over the next 11 years and took the brand from less than a 2% share to over 9%. Other Top 10 ads were used for years or set the scene for future executions.

An easy-to-follow clever storyline can be readily appreciated even in animatic or video storyboard format. Consumers can be confused by difficult storylines, or simple storylines in the rough that seem lackluster but can only be brought to life using finished production values. The opposite is true also. We have seen great ideas fail in market because of heavy-handed directing that overwhelms the storyline or because of changes in decisions about casting and the results of the rough testing.

Based on our research, we would suggest that that stimuli used for measuring or predicting woo should be as close to finished product as possible. If 9-year-olds can submit ideas created or co-created for YouTube, why can't professional ad agencies? Why is the industry still using animatics or video storyboards?

FINAL THOUGHTS ON THE WOO

The Beatles were wrong when they sang "can't buy me love". In business, love can be bought to improve a relationship with the right woo targeted at the right customer or consumer for a given brand in a specific culture. It is better earned but it can be bought.

Perhaps the science of genomics heralds the final frontier for woo. Knowing the DNA of customers or targeted consumers, their pre-programmed emotions, attitudes, motivations, and behavior may facilitate woo at the penultimate level.

All you need is love.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGMENTS

None declared.

REFERENCES

Dusenberry, P. (2005). *One great insight is worth a thousand good ideas.* New York: Portfolio.

Godin, S. (1999) *Permission marketing: Turning strangers into friends, and friends into customers*. New York: Simon and Schuster.

Hagerty, M. (1999). Testing Maslow's hierarchy of needs: National quality-of-life across time. Social Indicators Research, 46(3), 249-271.

Hill, D. (2005). *Body of truth.* Hoboken, NJ: Wiley and Sons, Inc.

Katsenelson, V. N. (2004, June 30). BJ’s wholesale vitaliy’s contrarian edge. Retrieved 12 May, 2010 from http://investmentinsight.blogspot.com/2004/06/bjs-wholesale.html

Moskowitz, H. R. & Gofman, A. (2007). *Selling Blue Elephants: How to Make Great Products that People Want Before They Even Know They Want Them.* Upper Saddle River, NJ: Wharton School Publishing.

Precourt, G. (Ed.) (2007). CMO thought leaders: The rise of the strategic marketer. McLean, VA: Booz Allen Hamilton.

Roberts, K. (2004). *Lovemarks: The future beyond brands.* New York: PowerHouse books.

TD Bank advertising and media room (2009). *Awards*. Retrieved 29 November, 2009, from https://mediaroom.tdbank.com/index.php?s=59

Wikipedia (2009). List of emotions. Retrieved 25 November, 2009, from http://en.wikipedia.org/wiki/List_of_emotions

Wordnet Search (2009). Need. Retrieved 30 November, 2009, from http://wordnetweb.princeton.edu/perl/webwn?s=need&sub=Search+WordNet&o2=&o0=1&o7=&o5=&o1=1&o6=&o4=&o3=&h=

PART V

STRUCTURED PACKAGE AND WEBSITE OPTIMIZATION

CHAPTER 15

Helping Packages Get Noticed on the Shelf Using RDE

Alex Gofman[1], Howard R. Moskowitz[1,*] and Tõnis Mets[2]

[1]Moskowitz Jacobs Inc., White Plains, New York, USA and [2]University of Tartu, Tartu, Estonia

Abstract: This chapter explores approaches to consumer-driven optimization of package design utilizing rule developing experimentation (RDE). The approach comprises dynamic creation and testing of a large number of design prototypes with consumers. RDE then uncovers optimal solutions, both on an aggregated, segmented (pattern-based latent mind-set segmentation) and on an individual-by-individual basis. Disciplined experimentation produces more targeted package designs, generating higher appeal to the consumers. The proposed steps describe a fast, parsimonious and actionable process, applying RDE to packages, providing in turn necessary input to designers about consumer preferences. The chapter demonstrates that systematic exploration using experimental design should be a central part of the initial, knowledge-gathering phase in package design. The steps of fitting the research into the package design process are shown. These steps constitute a cost-effective and efficient way to include consumers in the early stages of package design.

Keywords: Conjoint analysis, experimental design, interaction, package design, package design optimization, regression analysis, rule developing experimentation, suppression, synergism.

INTRODUCTION

Package design plays a critical role in purchase decisions. Approximately 73% of such decisions are made at the point of sale (Connolly and Davidson, 1996), where more attractive packaging frequently wins (Rettie and Brewer, 2000). Silayoi and Speece (2007) argued that when the consumer is undecided, the package becomes a critical factor in the purchase choice because it communicates to consumers at the decision-making time. Löfgren (2008) called this time "the first moment of truth" when the package functions as a silent salesman. Silayoi and Speece (2007) further suggested that *how* consumers perceive the subjective

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

entity of products, as presented through communication elements in the package, influences choice. In turn, this perception may be the key to success for many marketing strategies.

Underwood, Klein and Burke (2001) demonstrated that visuals on the package can be a strategic method to differentiate one's product. Pictures are more effective stimuli compared with words. Furthermore, consumers process visual information faster and easier, particularly in low situations where there is actually low involvement with the product. The correct selection of package colors plays a very important role as well. Grossman and Wisenblit (1999) and Madden, Hewett and Roth (2000) suggested that colors often create potentially strong associations for consumers, driving their brand preferences.

This chapter shows how structured experimentation in the form of rule developing experimentation (RDE) provides designers with a disciplined approach. RDE sifts through possible design features and their combinations. These designs, created out of the designer's talent, are then systematically explored to discover what works for a package design. RDE generates a narrower set of design options that are the most acceptable to consumers. This selection is based on "hard", statistically robust consumer reactions obtained early in the development cycle. The acquisition of this data from systematized stimulus arrays generates a more productive process for design. Productive guidance to artists/designers ensures that they can concentrate on the more profitable direction.

OPTIMIZING PACKAGING

For most of human history, packaging was utilitarian. Approximately 15,000 years ago, late Paleolithic settlers in Japan produced some types of pottery. It is quite conceivable that they or other early humans adorned their creations with the same fascinating images we find now on the walls of the caves they lived in (Arnold, 1985). Ancient civilizations witnessed some of the early known usage of art and graphics on food-related packages in the form of artistic amphorae, *etc.* albeit limited to upper classes.

The main purposes of food packaging were to provide a safe and convenient storage for the food, protect it from spoilage and pests and facilitate easy transportation. The aesthetic side of wrapping the mainstream food came only in the past 200 years (Klimchuk and Krasovec, 2006).

While serving the four main functions of packaging—containment, protection, convenience, communication (Robertson, 2005)—the technological marvels that keep, for example, milk unspoiled for years, were beautified by the top designers, making packaging into a commercial art form. This is clearly demonstrated in Saito (1999), who researched the issue in Japan. This chapter points out that the Japanese aesthetic traditions are deeply exemplified in the art of packaging. Unfortunately, the artistic approach also ends up misguided. A shopping trip to various stores will reveal creations of art on the shelves without regard to consumer needs and tastes, sustainability and environmental issues, *etc.*

Gomez (1999) suggested that the approaches that work well in other media, *e.g.,* minimalist design, do not apply to package design graphics. The products must figuratively jump out at the consumer in order for the consumer to select it from the shelf. While on the shelf, the product competes with many other offerings. Experts advise that the graphics designer should be as bold as package configuration, space and stacking position allow, using lively, persuasive colors, striking typefaces and prominent, creative photography or illustration (Jarman, 1999).

Multiple stakeholders with very different views and goals are involved in packaging—marketers, designers, product developers, brand managers, *etc.* Each stakeholder tries to improve the creation. All too often the results are disastrous. The packages look “too busy”, with too much graphic and text information, some of which might overwhelm, whereas other information may simply be irrelevant. Although one would think that there is little or no harm in placing an irrelevant message or a visual on the package, the desired consumer response, either comprehension or selection, might actually suffer consequently. There is evidence that the irrelevant information weakens consumers’ beliefs that the product will provide the benefit (Meyvis and Janiszewski, 2002).

It is difficult to overstate the role of correctly choosing the right visual parameters for packaging. Even when shoppers are open-minded and directly considering a category as opposed to picking up their usual brand, more than one-third of the brands displayed are completely ignored. However, a unique appearance consistently helps attract shoppers' considerations and drives purchase (Young, 2008).

One of the ways to design winning packages involves *experimentation.* Experimentation, argue Thomke (2001, 2003) and Kahn, Barczak and Moss (2006), lies at the very heart of new product development (NPD) and thus is connected to corporate values, habits, strategies and organizational structures. Thomke (2003) further pointed out that experimentation as an essential part of NPD and new product launches has enjoyed its attention only recently. Thomke (2001, 2003) also demonstrated the importance of doing the experiments quickly, quoting Edison by saying that a real measure of success is the number of experiments that can be conducted in 24 hours. Although it is an extreme opinion, it underscores the increased pressure of the competitive environment.

Perhaps the most dependable and effective way to satisfy consumers is to involve them in the actual process of creating the package. Focus groups and other forms of direct questioning on a post-hoc basis, although still popular, do not usually produce actionable results. The groups or survey techniques ask the consumer to evaluate *what has been created* and identify aspects of the package or product that are liked *versus* those that are disliked. The problems associated with the actionability of simple post-hoc evaluations have led to other approaches. For example, to increase the actionability of the consumer involvement some researchers and practitioners have gone so far as to abandon completely attempts to *understand* user needs in detail in favor of *transferring* need-related aspects of product and service development to users through use of so-called toolkits (von Hippel and Katz, 2002). This latter part is purely utilitarian, with the goal to create the product and service by an evolutionary approach that does not, however, produce knowledge of rules or reasons "why".

Consumers should *co-create* the package in one form or another in order to ensure that the consumer will eventually buy it. The full range of consumer involvement on every step of new package design creation is beyond the scope of this chapter (see, *e.g.,* Thomas, 2008). We concentrate on selecting the right package graphics and the involvement of consumers in both rule development and co-creation.

There has been substantial research done in using conjoint analysis for analyzing package design (Mohn, Roane and Stanton, 1982; Green and Srinivasan, 1990; Balakrishnan and Jacob, 1996; Rokka and Uusitalo, 2008). Silayoi and Speece (2007) presented a case of applying conjoint analysis for researching the importance of package attributes. The approach uses a limited number of experimentally designed prototypes that are manually pre-created for the study. The effort results in a number of interlinked statistical problems in the analytic phase identified in Part 1.

APPLICATION OF RDE TO IDENTIFY WHAT WORKS IN PACKAGE GRAPHICS

This chapter demonstrates that RDE applied to graphics design is very similar to RDE with just words or with text and pictures (see also Gofman and Moskowitz, 2009).

The case study demonstrates this approach using the example of a package for shampoo, both for the total panel and for latent segments that exist in across the array of consumers. Rather than create a single package for evaluation, the designer creates a template with various features (*e.g.,* picture of the product) and several options for each specific feature in the template. The goal then is to identify what drives consumer interest and what does not.

RDE requires the creation of multiple packages based on an experimental design. In our example, we divide the target package into six features (Fig. **1**): caps (color variations), brand name, shampoo type, name, main picture and health message, respectively. All of the alternative executions of a specific feature should match each other and fit the outlining package (in the center of Fig. **1**).

The design in Fig. **1** shows the package “recipe”, *i.e.* the combination of options for each test package. The design mixes and matches one option of each feature to create what can essentially be considered a prototype design. There will be many of recipes or alternative prototypes to evaluate. The respondent need only evaluate the different combinations, without necessarily realizing that the different combinations have been created according to a structured approach.

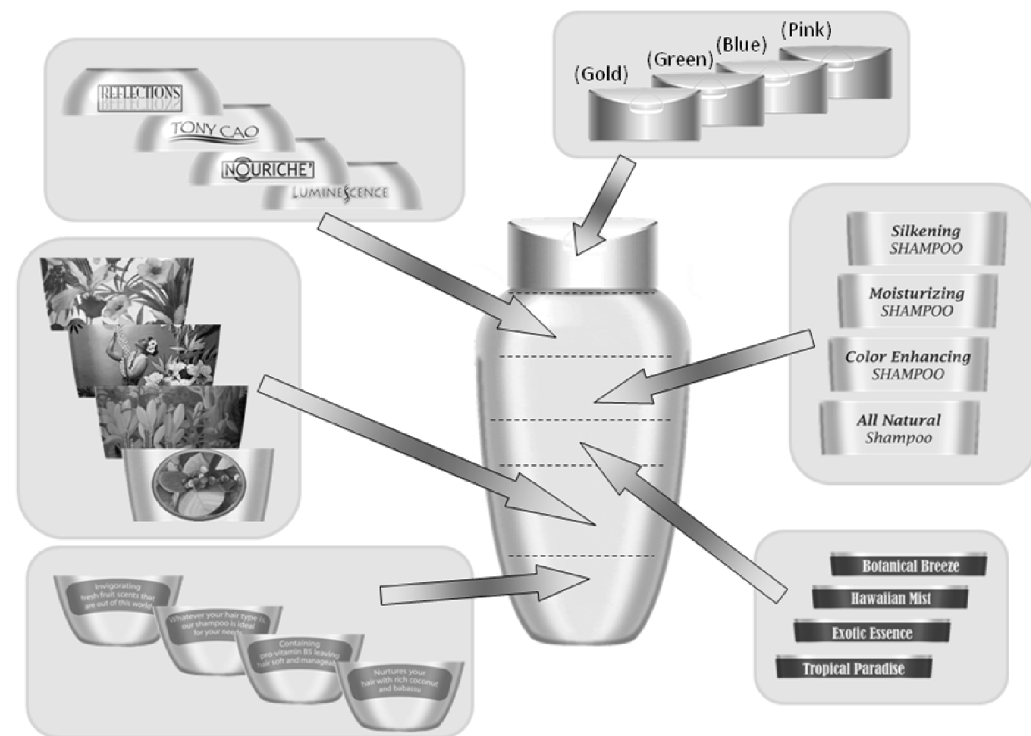


Figure 1: The background layer (center) and the features of the package.

The experimental design prescribes the different recipes for the package. The graphical elements (options) are combined/overlaid one with another in order to create the specific renderings of the prototypes that will be tested on the computer screen. The challenge is that, unlike a box of, say, cereal, a shampoo package is curved, which makes it more difficult to mix and match the visuals into a prototype that looks realistic or at least coherent and seamless. *The obvious question was how to work on a two-dimensional screen with a three-dimensional problem. The solution is surprisingly straightforward.*

Each feature of the package can be thought of as a transparent layer. In a way similar to that done by Adobe Photoshop®, the features of the project are transparent everywhere except for the key object of the layer. For example, Fig. **2** shows six layers with individual features located in each layer. In the actual layers, the area outside a particular component would be made completely transparent.

Figure 2: "Building" the experimental packages as layers (the outlines of the package on the layers are for demonstration purposes only).

The designers first created several executions of each of the six features of the shampoo package (see Fig. **1**). Each option was created on a transparent layer. The feature was positioned properly to allow correct matching during process of overlaying one transparency atop another. All the surface bending was captured inside each option. The designer can do that quite easily based on a single template, shown in the middle of Fig. **1**. The result of the layering is a graphically realistic picture, assembled by the principle of RDE. The experimental design prescribes these different pictures.

The computer (browser) superimposes these transparencies according to the RDE design, thus creating different executions of the experimental packages. Each new combination defined by the RDE design corresponds to a different package. During the course of the interview, the participant evaluates many different combinations of options. The participant never sees the individual transparencies, but only complete packages (see Fig. **3** below). The individual transparencies are already in place on the participant's computer because the server uploaded them at the start of the interview, a stratagem that speeds up the screen changes and creates in its wake a more pleasant experience for the respondent (Moskowitz *et al.* 2004).

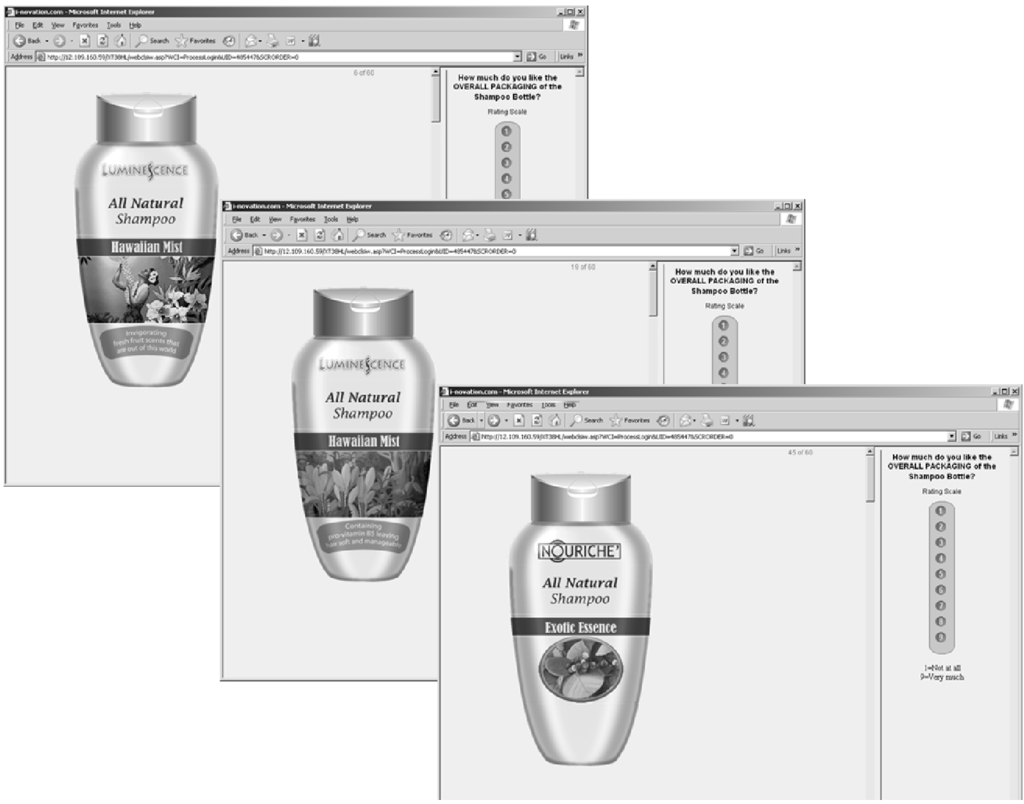


Figure 3: Sample experimental packages. The bottom package has a missing option.

When participants are exposed to these synthesized packages, they do not really know that the experimental design lies underneath the combinations, nor could they. The transparencies are combined so quickly that, to the participant, it looks like a single package. The participants evaluate package after package, one at a time. When the evaluations last only approximately 12 to 18 minutes, most respondents have no problem and actually say that they enjoy the experience. This is because people assess visual stimuli much faster than it takes them to read text. This speed of the response, almost a gut-feel response, compensates for the increased number of concepts that graphical RDE uses.

RDE uses experimental designs that, on occasion, call for the absence of some components in some test concepts (prototypes). The in-depth analysis of the issue is beyond the scope of this chapter. Suffice it to say that the absence of a component in the test concepts allows the regression analysis to generate absolute values for the estimated utilities. What concerns us here are the emotional reactions. According to Gofman and Moskowitz (2005), despite the designer's "concern" that the respondents will be taken aback by incomplete designs that are lacking something, the reality is just the opposite. The design does not have to be complete. Participants have no problem evaluating both the complete and the partial packages (Fig. **3**, bottom screen). When asked after the interview through an exit survey whether they felt uncomfortable, almost no participants reported feeling uncomfortable with partial, incomplete packages. Indeed, most respondents didn't care; they evaluated what was in front of them. The care and concern is more often voiced by the design professional, not by the consumer. And the concern is intellectual, rather than data based.

A very important difference from a text-based RDE is the need for a filler background image for those situations when the package on the screen occasionally must be without an option, as required by the experimental design. The best solution places a bare package in the back of each screen, behind all the layers. This way, a zero condition (the absence of an option in the design) does not create a disturbing image with "holes" in it. The background image of the bare package makes the test stimulus on the computer screen look like an acceptable package, meaningful to consumers even if it has an element missing.

This design creation and research exercise merges art, science and consumer knowledge formalized by RDE, easily and timely executed *via* the Internet. The consumer participants report on exit interviews that they have evaluated realistic-looking packages, rather than rough designs, ensuring a modicum of reality.

ANALYSIS

Creating the Model

Experimental design in RDE begins by creating a unique balanced design for each individual respondent. Each respondent evaluates different combinations, albeit with the same set of elements. The design of one respondent is isomorphic to that of another. Just the combinations are different. The RDE data allow for estimating an equation for an individual respondent that relates the presence/absence of each of the design feature to the rating assigned by the respondent.

The model thus shows how every package element drives the response (*e.g.,* a rating "1 = Not interested → 9 = Very interested"). The individual modeling, done by regression analysis, will generate understanding of the package features and help to design packages that are more impactful.

The key to understanding the data for product development and innovation is the impact of each individual element. To the degree that the elements perform well the developer will be able to synthesize new and potentially breakthrough shampoo package concepts. To the degree that the concept elements perform only modestly, the concepts will not be breakthrough. The norms for interpreting utility scores and additive constants are listed in Table **1**. Here the numbers are conditional probability (percentage) of people being interested in buying the product as shown by the picture (*i.e.* the visual test concept). The additive constant in RDE modeling with the incomplete concepts is a baseline interest of respondents in the package the product (without any elements present).

Table 1: Norms for the Additive Constant and the Utilities (Moskowitz *et al.* 2005).

Additive	**Interpretation**
Constant	
>60	Respondents are very predisposed to the product
50–60	About half of the respondents are very positive
40–50	Respondents accept the idea
30–40	The elements need to do the work
<30	The product is a commodity and the elements must do the work
Utility score	
>15	The element performs exceptionally well, breaks through clutter
10–15	The element performs well
5–10	The element breaks through the clutter
0–5	The element is barely effective, if at all
<0	The element actively detracts from acceptance

Table **2** shows the type of data about the elements of package design that emerge from the exercise. Recall that each respondent generated a set of impact numbers, or coefficients from the regression analysis. The respondent ratings for the different package designs were originally assigned on a 9-point scale. However, following the conventions of market research, the ratings were re-coded. Ratings of 1–6, the lower part of the scale, were re-coded to 0 to denote that the respondent looking at the particular package design was not interested, or at best marginally interested. In contrast, ratings of 7–9, the higher part of the scale, were re-coded to 100 to denote that the respondent was interested. Afterwards, the RDE tool ran the regression analysis, relating the presence/absence of the design features to the binary response of "not interested" or "interested". Table **2** shows the parameters of the regression for the total panel and for four emerging pattern-based mind-set segments with different points of view.

Table 2: Performance of Options for the Shampoo Packaging Study (Total and Two Segments).

	Total	**Seg1**	**Seg2**	**Seg3**	**Seg4**	**Males**	**Females**
Base Size	183	75	33	50	25	20	163
Constant	3	–3	–1	8	18	10	2
Cap Color							
A1 Pink	0	–3	–4	4	5	2	–1
A2 Blue	0	–2	–3	3	4	0	0
A3 Green	–1	–2	***–6***	1	4	1	–1
A4 Yellow	–1	–3	–4	1	4	2	–1
Brand							
B1 Nouriche	0	–4	**10**	–3	4	2	0
B2 TonyCao	1	–4	**12**	–2	3	3	0
B3 Reflections	2	0	**11**	–1	5	2	2
B4 LumineScence	1	–3	**11**	–1	3	0	1
Shampoo Type							
C1 All Natural	5	0	**15**	4	5	2	5
C2 Silkening	5	3	**16**	5	–4	–1	**6**
C3 Moisturizing	6	2	**21**	**6**	–2	0	**7**
C4 Color Enhancing	3	–4	**25**	2	–1	***–10***	5
Name							
D1 Hawaiian Mist	1	5	1	1	***–8***	–2	2
D2 Exotic Essence	2	4	0	4	***–7***	2	2
D3 Tropical Paradise	2	4	1	1	***–6***	4	1
D4 Botanical Breeze	1	4	–1	2	***–6***	3	1
Main Picture							
E1 Flowers	**18**	**19**	**19**	**28**	–5	–1	**20**
E2 Woman	**14**	**13**	**14**	**26**	***–8***	2	**15**
E3 Flower Circle	**16**	**17**	**13**	**25**	–3	3	**17**

Table 2: cont….

E4 Blue Forest	**14**	**17**	**14**	**20**	–5	4	**16**
Health Message							
F1 Nurture Coconut	**7**	**12**	3	1	**9**	4	**7**
F2 Fresh Fruit	**9**	**19**	4	1	5	5	**10**
F3 Vitamin B5	**7**	**15**	2	1	4	3	**8**
F4 Whatever Hair Type	**7**	**13**	1	1	**7**	4	**7**

Note. Numbers in the body of the table are the impact values, after the ratings have been converted to a binary scale (ratings 1–6 → 0; ratings 7–9 → 100). The bold values add substantially to the liking; the bold italic detract from it.

Quite often when we deal with text concepts about a product, such as shampoo, we will see high additive constants. The additive constant is the expected rating (*e.g.,* on the binary scale, 0/100) for the case of a text concept without any elements, *i.e.* the idea of a shampoo, but no elements. We see these high constants because people like the idea of shampoo as a product, even without description. When we deal with packages, however, people are focused on exactly what they see. In the case of packages of shampoo without any elements, we see nothing. There is no basic interest in a package of shampoo with nothing. And hence we have low additive constants, near 0 and sometimes lower.

Analyzing the Data: What's Working; What's Not

As in any other RDE study, the essence of the study is that one table that reveals the impact or utility of every element. RDE generates these data. Even without paragraphs of interpretation, simply scanning the results quickly reveals what's going on in the mind of the respondent. The verbiage surrounding the analysis simply clarifies the patterns that the numbers suggest. With that in mind let us take a look at the results. It's generally easiest to enumerate the results, each as a specific finding and afterwards pull together the different themes.

Looking at Table **2**, we should be struck by three things: (1) the big numbers, *i.e.* the most impactful elements cluster into the three silos of picture (biggest impact), then (2) the health message and then (3) shampoo types.

Going further into the data, we could pull out the different scoring elements in each silo. What should further attract our attention is that it is not the silo or category but rather the specific element. We cannot, nor should we ever, draw a blanket conclusion such as "picture is important for packages" without looking at the impacts of the different elements. The pattern and conclusions are in the details. Only by exploring many different elements in a single silo can we make the conclusion:

1. RDE forces all of the elements to appear equally often.

2. There is no necessary reason for a silo or category to comprise elements that all perform equally. The elements are independent agents.

3. So when we see that pictures all score high but differently from each other, we then can conclude that pictures are most important for the study. Yet the fact that the pictures perform differently should tell us right away that the respondent not only pays attention to the pictures but also clearly differentiates among them.

4. We don't see it here, but there is every possibility in an RDE study that some pictures might perform terribly. We simply did not work with those poor performers. In fact, in many RDEs, pictures that are poor performers are not used. The graphic elements are often developed by designers, or by amateurs who are on the lookout for poor performers ahead of time. These poor performing graphic elements are often obvious and edited out. In contrast, when RDE is run with text elements rather than graphic elements, the editing process up-front is not quite as strong, or perhaps better stated, the poor quality of the text element is not quite as obvious as the poor quality of the graphic element.

5. For the total panel, the most important features, *i.e.* those with the highest impact, are the main picture and, to a lesser extent, the health message.

6. The impact values are very high for the Main Picture category (from +14 for the picture of a woman to +18 for the picture of flowers).

7. The health messages have impacts of +7 to +9. Although positive and important in increasing interest, they are not different enough to cause variation in impacts.

8. The same (to a much lesser extent) applies to Shampoo Types (+3 to +6, which is considered neutral, with a slight positive skew).

9. Cap color, brands and names have no impact on the total panel (all the impacts fall in the neutral zone).

10. Here are notable differences in impacts of individual elements for males and females subgroups (Table **2**). Although caps, brands and names are neutral for both subgroups, they greatly differ in their reaction for other categories:

a. Main pictures are highly impactful on females (+15 for the picture of a woman to +20 for the flowers, while keeping men indifferent (–1 to +4).

b. Health messages are quite positive for women (+7 for *Nurture Coconut* and *Whatever Hair Type* to +10 for *Fresh Fruits*), while still neutral (although in a slight positive territory) for men.

c. Men are generally neutral to shampoo types with the exception for *Color Enhancing* (–10), whereas women are neutral to slightly positive to the ideas with moderately good reaction to *Moisturizing* type.

To help better understand the results, the utilities could be mapped graphically. Fig. **4** shows the results of RDE analysis for the female subgroup mapped in graphical form on a chart. The higher the position of the element is on the chart, the more positive impact it has on the purchase intent. If the elements are located around zero (±5), their impact is neutral. The elements plotted below –5 detract from the purchase intent. The horizontal placement of elements is not meaningful and is dictated by available space. The ovals in the left part of the chart represent the range of the utilities for each category (from the lowest to the highest in that category). For example, categories A (cap colors), B (brand) and D (name) demonstrate very little variation in the utilities of their elements (all neutral), whereas category F (health message) and E (picture) show the biggest range of utilities.

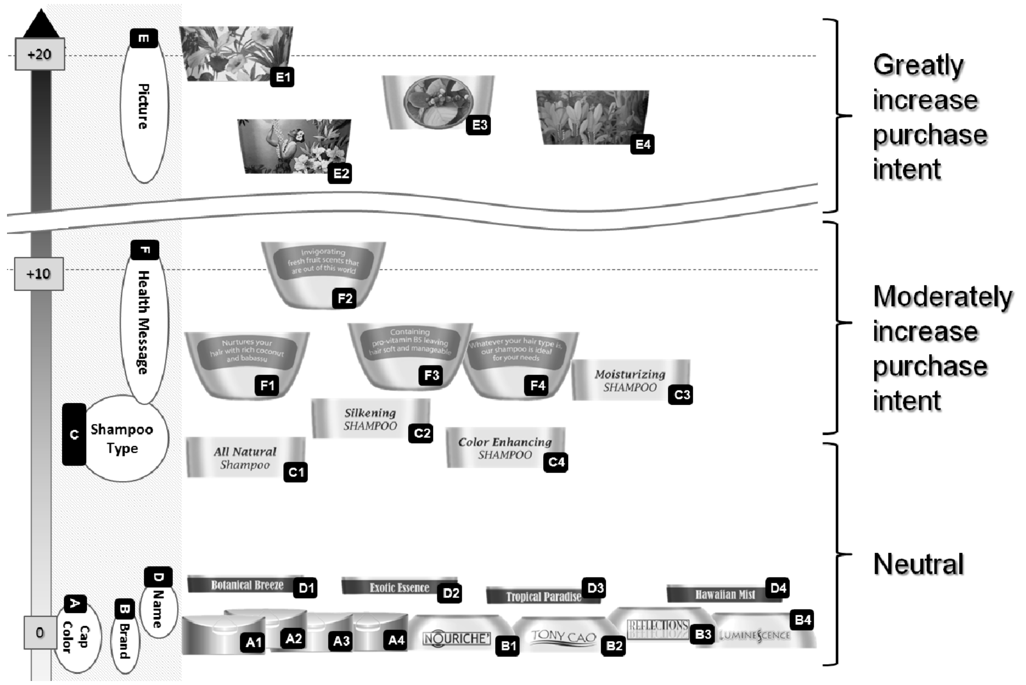


Figure 4: Performance of individual elements for female consumers. The numbers are the conditional probability that consumers would be interested in purchasing the product if a particular element is present. The vertical location of the elements represents their utilities values. Ovals (on the left) show the range of the impacts for the category.

The most impactful (optimized) shampoo packages for the total sample and the male subgroup are shown on Fig. **5** (left and right packages respectively). The packages are synthesized based on the highest ranked impacts from each category (Moskowitz, Porretta and Silcher, 2005).

Consumers Differ in a Profound Way Regarding Packages—Mind-Set Segmentation

Consumers are not all created the same. People's preferences differ substantially. Most traditional approaches divide people based on some demographic or purchase behavior criteria. A more effective approach for design and development divides people based on their mind-sets,and then specifies the development accordingly. People in the same mind-set segment like the same package design features or products. We can cluster respondents based on the patterns of their responses and try to optimize the package for each segment. This approach has proved itself to be quite robust in many dozens of case histories (Gofman and Moskowitz, 2005).

Figure 5: Optimized packages. The optimal packages: for Total panel (left), for Segment 2 (middle) and for males (right).

The segmentation analysis revealed four substantially different mind-sets of the consumers:

Segment 1: Health-Oriented (roughly 40% of the total sample)

1. The majority of the consumers belong to this segment and they respond positively to health messages (from +12 for *Nurture Coconut* to +19 for *Fresh Fruits*) as well responding positively to those main pictures that could be associated with health (from +13 for the picture of a woman to +19 for flowers).

2. Segment 1 respondents are neutral to the remaining the elements with a slightly more positive attitude to the names and more negative (yet still in the neutral territory) to the cap colors and brands.

3. All-in-all, the health-oriented segment cares only about how healthy the shampoo is. Segment 1 prefers striking pictures.

Segment 2: Function and Image (roughly one-fifth of the total sample)

1. The consumers in this segment react more positively to more features.

2. Similarly to Segment 1, they react highly positively to the main picture (+13 to +19).

3. Specific shampoo types drive their purchase intent even more, reaching an extraordinarily high +25 for *Color Enhancing.*

4. Segment 2 is very positive to the brands although they don't differentiate among the options (+10 to +12).

5. Segment 2 is neutral to the rest of the features. They slightly dislike the colors of all the colors of the caps (–3 to –6).

6. They like something that stands out on the shelf.

7. For a designer to create an appealing package for Segment 2, it is quite important to know both positively accepted (by the consumers) elements of the packaging as well as negative ones. The optimized package for Segment 2 appears in the middle of Fig. **5**.

Segment 3: Visual (roughly one-quarter of the total sample)

1. This segment reacts strongly positively to the main picture (+20 to + 28) while being mostly neutral to everything else.

2. All-in-all, this segment wants a striking image on the package. Everything else is irrelevant to them.

Segment 4: Skeptics (roughly a one-sixth of the total sample)

1. Nothing excites this segment.

2. It is marginally positive to the health messages (+9 for the *Nurture Coconut*) while being neutral to a slight negative reaction to the main picture.

In a typical case, or better in a perfect world where shelf space is not at a premium, a manufacturer might consider creating a separate SKU (stock keeping unit) for each segment. The separate SKUs will maximize the appeals to the different mind-sets of the consumers. Otherwise, the manufacturer will still do well to create a single SKU optimized for the total panel. In that case, although capturing fewer buyers than with individually optimized packages for each segment, the manufacturer will create a more appealing product than a package created on the basis of a simple guess.

SUMMARY AND PRACTICAL IMPLICATIONS

The proposed extension of RDE to graphical elements applies conjoint analysis to uncover consumer preferences of packages. The metaphor of transparent layers superimposed upon each other and the use of the statistical power of individual permuted experimental designs define both the necessary combinations and how to use these combinations in a "battle-tested" manner, with realistic-looking designs. The approach solves some of the problems for a case where the consumer knows what he wants but cannot articulate it. At the same time, the design approach overcomes the complex and intervened statistical problems related to traditional methods of conjoint analysis utilizing a single experimental design for a project.

The application of RDE to graphical optimization provides important and timely input about consumer preferences to designers. RDE does not replace the artistic talent in any way. Quite to the contrary, RDE narrows the multiple choices available for the designer, leading to what aspects affect the consumer respondent. RDE is, in the end for designers, a cost-efficient mechanism to aid creation, finding use in those mature categories where differentiation is difficult and the flexibility for really new development is small. RDE becomes part and parcel of a company's intellectual property (IP). The information, the structure and segments become integral to understanding opportunities and developing future products (Gofman, Moskowitz and Mets, 2009a, 2009b).

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This chapter is partially based on Gofman, A., Moskowitz, H. and Mets, T. (2010). Accelerating structured consumer-driven package design. *The Journal of Consumer Marketing, 27*(2), 157–168.

REFERENCES

Arnold, D. (1985). *Ceramic theory and cultural process.* Cambridge: Cambridge University Press.

Balakrishnan, P. V. & Jacob, V. S. (1996). Genetic algorithms for product design. *Management Science, 42*(8), 1105–1117.

Connolly, A. & Davidson, L. (1996). How does design affect decisions at point of sale? *Journal of Brand Management*, 4(2), 100–107.

Gofman, A. & Moskowitz, H. (2005). Developing from the ground up: Self-authoring systems for text and package concepts in Concept Research. In H. R. Moskowitz, S. Porretta, & M. Silcher (Eds.), *Food product design and development* (pp. 283–322). Ames, IA: Blackwell Publishing.

Gofman, A. & Moskowitz, H. (2009). *Structured experimentation in package design*. Wharton School Publishing. Retrieved April 27, 2009, from http://www.whartonsp.com/articles/article.asp?p=1334083&seqNum=5

Gofman, A., Moskowitz, H., & Mets, T. (2009a). Integrating science into web design: Consumer driven website optimization. *The Journal of Consumer Marketing, 26*(4), 286–298.

Gofman, A., Moskowitz, H., & Mets, T. (2009b). Developing new corporate understanding of an existing product. *The Journal of Product & Brand Management, 18*(2): 84-94.

Gomez, E. M. (1999). *New design: Tokyo*. Beverly, MA: Rockport Publishers.

Green, P. & Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing, 54,* 3–19.

Grossman, R. & Wisenblit, J. (1999). What we know about consumers' colour choices. *Journal of Marketing Practice: Applied Marketing Science*, *5*(3), 78–88.

von Hippel, E. & Katz, R. (2002). Shifting innovation to users *via* toolkits. *Management Science, 48*(7), 821–833.

Jarman, J. B., Jr. (1999). How to take your package design to the next level. *Frozen Food Age, 48*(4), 22.

Kahn, K. B., Barczak, G., & Moss, R. (2006). PERSPECTIVE: Establishing an NPD best practices framework. *Journal of Product Innovation Management*, *23*(2), 106–116.

Löfgren, M. (2008). Customer satisfaction in the first and second moments of truth. *The Journal of Product and Brand Management*, *17*(7), 463–474.

Madden, J., Hewett, K., & Roth, M. (2000). Managing images in different cultures: A cross-national study of colour meaning and preferences. *Journal of International Marketing, 8*(4), 90–107.

Meyvis, T. & Janiszewski, C. (2002). Consumers' beliefs about product benefits: The effect of obviously irrelevant product information. *The Journal of Consumer Research, 28*(4), 618–635.

Mohn, N. C., Roane, S. B., & Stanton, W. W. (1982). Forecasting consumer preferences using conjoint measurement. *European Research, 10*(3), 106–112.

Moskowitz, H., Gofman, A., Manchaiah, M., Ma, Z., & Katz, R. (2004). Dynamic package design and optimization in the Internet era. In *Proceedings of ESOMAR Technovate 2 Conference: Worldwide technology and innovation conference* in Barcelona, February 1-3, 2004 (Ed. Fellows D.), 173–218. Amsterdam: ESOMAR.

Moskowitz, H., Porretta, S., & Silcher, M. (2005). *Product design and development.* Ames, IA: Blackwell Publishing.

Rettie, R. & Brewer, C. (2000). The verbal and visual components of package design. *Journal of Product & Brand Management, 9*(1), 56–70.

Robertson, G. L. (2005). *Food packaging: Principles and practice* (2nd ed.), Food Science and Technology Series, Vol. 152. Boca Raton, FL: CRC Press.

Rokka, J. & Uusitalo, L. (2008). Preference for green packaging in consumer product choices—Do consumers care? *International Journal of Consumer Studies, 32*(5), 516–525.

Saito, Y. (1999). Japanese aesthetics of packaging. *The Journal of Aesthetics and Art Criticism, 57*(2), 257–265.

Silayoi, P. & Speece, M. (2007). The importance of packaging attributes: A conjoint analysis approach. *European Journal of Marketing, 41*(11/12), 1495–1517.

Thomas, J. (2008). Begin at the beginning: Research should be involved at the outset of any (re)packaging process, *QUIRK'S Marketing Research Review, XXII*(1), 52–56.

Thomke, S. (2001). Enlightened experimentation. The new imperative for innovation. *Harvard Business Review, 79*(2), 66–75.

Thomke, S. (2003). *Experimentation matters: Unlocking the potential of new technologies for innovation*. Boston, MA: Harvard Business School Press.

Young, S. (2008). New and improved indeed: Documenting the business value of packaging innovation. *QUIRK'S Marketing Research Review, XXII*(1), 46–50.

Underwood, R. L., Klein, N. M., & Burke, R. R. (2001). Packaging communication: Attentional effects of product imagery. *Journal of Product & Brand Management, 10*(7), 403–422.

CHAPTER 16

Introduction to Consumer-Driven Optimization of Landing Pages

Alex Gofman*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: Although widely recognized as important to increase a website's conversion rate and overall ROI, landing page optimization (LPO) remained for a long time in the domain of IT. This chapter shows the development, classification, advantages and shortcomings of a very advanced form of LPO, multivariate landing page optimization. The RDE approach tests thousands of Web page prototypes with consumers and finds real optimal solutions on an aggregated, segmented and individual basis. The latter paves the road to individually optimized pages and one-on-one marketing.

Keywords: Conjoint analysis, landing page optimization, multivariate analysis, rule developing experimentation.

INTRODUCTION

The issue of understanding consumer preferences is critically important for innovation in different stages of product development and marketing (Drucker, 1995, 2002; Von Hippel, 2005; Von Hippel and Katz, 2002). Corporations realize that one of the ways to build consumers' trust and interest in their products and services, ultimately leading to improved conversion rates, is consumer-centric Website design (Berland *et al.* 2001; Chandler and Hyatt, 2002; Palmer, 2003; Schlosser, White and Lloyd, 2006). Experimentation with consumers is one of the most powerful approaches used to obtain actionable consumer insights and achieve such goals (Janssen and Dankbaar, 2008; Gofman, Bevolo and Moskowitz, 2009).

During the initial years of Internet, the only people involved in the creation of Websites were IT professionals and geeks. The stress was on "programming"—overstuffing the pages with a limited selection of "bells and whistles" afforded by the then-current version of HTML. In many cases, this resulted in sites that were painful to use (such as the blinking words and crawling "ants" of early 1990s).

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

The next stage was characterized by commercial artists accustomed to creating brochures and flyers, who pushed the pendulum toward the opposite direction. This period is remembered by very colorful, slowly loading artistic pages without much regard to usability. The third stage saw the entry of far more professional and focused Web designers, armed with relevant ergonomics rules to improve the reader's experience.

Usability labs addressed the problem of preventing ruinous designs and they worked to reduce and ideally to avoid disasters entirely, by approaching the problem from the "lower" end. Their focus was simple: how to avoid bad designs. The state-of-the-art thinking had not reached the stage of finding approaches that focused on synthesizing better, even optimal Web pages for specific audiences. It is here that we enter, to show how rule developing experimentation (RDE) in one of its instantiation—design of visual stimuli—provides a new opportunity for Web design, beginning with test stimuli, moving on into experimental design, then understanding and finally mind-typing to optimize the experience for the specific individual landing on a Web page.

The average bounce rate (percentage of single-page-view visits) on a Website is about 37% (White, 2006). Although some sites have a rate well above 50%, there are many others whose conversion rate lies in low single digits or even a fraction of 1%. A more serious problem according to some sources is the so-called *derelict conversion*. According to data from MarketingSherpa, the average e-commerce shopping cart has about a 60% abandonment rate. More graphically, this is equal to three out of five shopping carts in a department store abandoned in the aisles (Booth, 2006).

When Forrester Research evaluated the Websites of many major brands, the sites often failed even the most basic tests of usability and brand building, exhibiting failure rates of between 50% and 83% (Temkin, 2007). Today's skeptical and empowered customers have increasingly more access to information. They tolerate advertising far less. They are becoming ever harder to win and keep. Firms must raise the bar on the customer experience that their Website provides.

One of the culprits of the lackluster customer experience is self-centered design of the Websites. Companies often lack a sharp, research-based understanding of their target customers. Without this information, decision makers advocate experiences and features that they personally like. When an executive says, "I don't like this", or "That works for me", they typically focus on their own needs. Yet when the firm's target customers are teenage males, does it really matter how the experience feels to a 40-something female vice president of marketing (Temkin, 2007)? Unfortunately, all too frequently corporate dynamics wins. The highest-paid-person opinion (HiPPO) wins over hard consumer data.

LANDING PAGE OPTIMIZATION

For a long time, the only solution to improve the aesthetics of a Website was based on the subjective predilections of Web designers. This dependence on individual preferences, extended to the Internet audience, is prone to mistakes. Individual preferences do not anticipate the unknown. They cannot. People's preferences differ. Not taking these preferences into account may result in a loss of business. The potential loss of not optimizing the landing pages may be staggering. Furthermore, many Website designers do not consider the aesthetics of payment pages as being important. However, simple changes to those pages could bring a substantial improvement to revenue per visitor with some reporting boosting conversion rates as much as 600% (http://www.web-site-evaluations.com).

Researchers in Canada reported that the snap decisions Internet users make about the quality of a Web page have a lasting impact on their opinions. Impressions were made in the first 50 milliseconds of viewing (Lindgaard *et al.* 2006). These findings suggest that the main features and the general appearance of the landing page may well make a difference and not necessarily the actual content.

In the past few years, an approach called landing page optimization (LPO) became prevalent (see classification of the approaches at Fig. **1**). According to Gofman, Moskowitz and Mets (2009), LPO may either be target based (customization of the pages based on some known behavioral or self-profiled information about the visitor, *e.g.,* previous purchase record) or be experiment based (optimization of the pages based on the consumer's preferences obtained though some sort of experimentation). We will present the latter approach (experiment based) because it the most efficient and widely used.

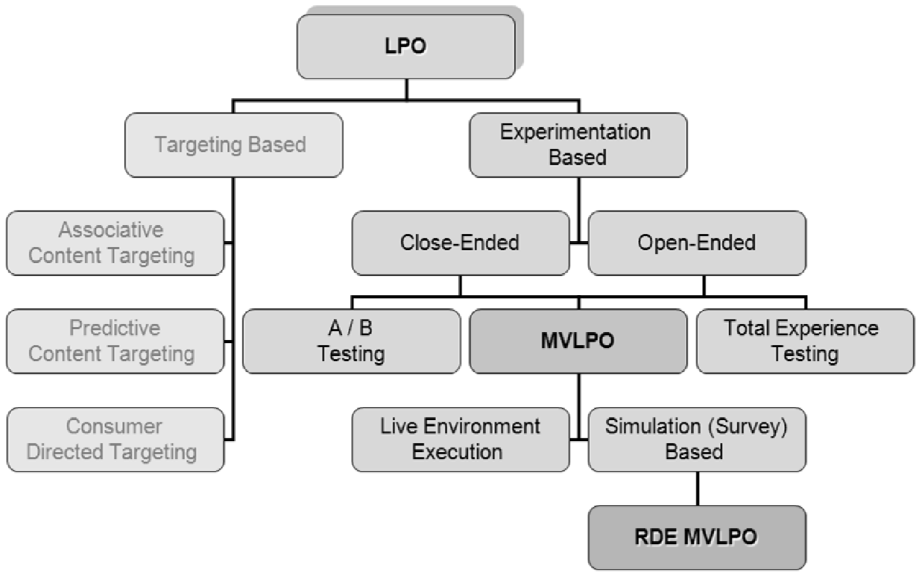


Figure 1: Classification of methods for landing page optimizations (LPO). Source: Gofman, Moskowitz and Mets (2009).

Three major types of LPO are based on targeting:

1. **Associative content** targeting (also called rule-based optimization or passive targeting). The page content is modified based on information obtained about the visitor's search criteria, geographic information of source traffic, or other known generic parameters that can be used for explicit non-research-based consumer segmentation.

2. **Predictive content targeting** (also called active targeting). The page content is adjusted by correlating any known information about the visitor (*e.g.,* prior purchase behavior, personal demographic information, browsing patterns, *etc.*) to anticipate (desired) future actions based on predictive analytics.

3. **Consumer directed targeting** (also called social targeting). The page content is created using the relevance of publicly available information through a mechanism based on reviews, ratings, tagging, referrals, *etc.*

There are two major types of LPO based on experimentation:

1. **Closed-ended experimentation** (limited in time). Consumers are exposed to several variations of landing pages and their behavior in response is observed. At the conclusion of the experiment, an optimal page is selected based on the outcome of the experiment.

2. **Open-ended experimentation** (ongoing). This approach is similar to closed-ended experimentation, except that the experimentation is ongoing. The landing page is adjusted dynamically as the experiment results change.

Traditional research measures customer responses, generally providing diagnostics, but no "rules" other than the intuitive guidance one gets from looking at results. For example, A/B split testing evaluates the differences in the visitor's reactions between a limited set of page executions (frequently, just two: A and B). Since it is just a "beauty contest", the visitor can only select the best from the tested solutions, which might not be the best possible (optimized) page. Traditional approaches evaluate only a few pre-selected executions of the pages (as in A/B split test) or the elements (*e.g.,* visuals) individually. Traditional methods do not test multiple options; generally, they do not develop "formal rules" to create optimal pages, although the data are often mined for so-called insights.

FROM TESTING TO RULES TO ACTION—MULTIVARIATE LANDING PAGE OPTIMIZATION

An advanced form of LPO, multivariate landing page optimization (MVLPO), involves hundreds or even thousands of prototypes. Introduced initially in late 1990s, MVLPO did not get the attention it deserved until very recently, especially with the launch of the Google Website Optimizer (Ash, 2008). A typical MVLPO creates multiple experimentally designed variations of a Web page and evaluates the difference in the reaction or behavior of the people who visit these pages.

LPO may focus on a single page (more typical) or the entire Website experience (newly emerging efforts). For example, total experience testing is an evolving approach that tries to optimize the whole Website experience rather than individual pages (Kaushik, 2007). Total experience testing is an extension of current thinking, which holds that it is the entire person–product or person–environment (situation) that is important, rather than the focused analysis of one small portion of that environment.

LPO can be executed in two different modes: live environment (experimentation is done with the production Website and regular visitors) and simulation/survey based (closer to traditional marketing research approaches with qualified respondents). The latter, frequently employing the power of RDE, has an advantage of not risking the alienation of valuable customers of the site with possibly suboptimal versions of the pages.

A typical MVLPO involves multiple experimentally designed variations of a Web page. The experiment creates a particular combination of multiple groups of elements (graphics, text, *etc.*) on a specific page. Each group comprises multiple executions (options). For example, a landing page may have *n* different options of the title, *m* variations of the featured picture, *k* options of the company logo, *etc.* An experimental design is applied to the elements of the page and the resulting prototypes are served to customers.

MVLPO constitutes a comprehensive, scientific approach to understanding customers' minds and using it to optimize their experience. MVLPO works on the basis of stimulus–response, creating the stimuli, testing, identifying patterns and creating equations, which allow for prediction and improvement. MVLPO evolved into an easy-to-use approach in which not much programming and IT configuration are needed. In many cases, a few lines of JavaScript on the page allow the remote servers of the vendors to control the changes, collect the data and analyze the results. MVLPO provides a foundation for a continuous learning experience.

At the same time, MVLPO may generate suboptimal results when the original materials are not chosen carefully [the famous, so-called GIGO ("garbage in, garbage out") effect]. Another limitation is that MVLPO usually optimizes one page at a time. Yet, the reality is that Website experiences for most sites are complex multi-page affairs. For a typical e-commerce Website, a successful purchase involves visiting around 12 to 18 pages. The support site may be even longer. For the holistic experience optimization, the total experience optimization approach could be considered (Kaushik, 2007).

EXECUTING MVLPO—THE PRACTICALITIES

MVLPO can be executed in a live (production) environment (*e.g.,* Google Website Optimizer, Optimost.com, *etc.*) or through a market research survey/simulation (*e.g.,* Moskowitz Jacobs, Inc.'s StyleMap®.NET).

In *live environment MVLPO execution*, a special tool (server) makes dynamic changes to the Website. The visitors are directed to different executions of landing pages created according to an experimental design. The system keeps track of the visitors and their behavior (including their conversion rate, time spent on the page, *etc.*) and with sufficient data accumulated, estimates the impact of individual components on the target measurement (*e.g.,* conversion rate).

With an adequate number of observations, this live environment approach is very reliable because it tests the effect of variations as a real-life experience, generally transparent to the visitors. Live environment evolves toward a relatively simple and inexpensive approach (it applies to Google Optimizer at the time of this writing). In contrast, it may take a long time to achieve statistical reliability caused by variations in the amount of traffic, which generates the data necessary for the decision. The live environment approach may not be suitable for low-traffic/high-importance Websites when the site operators do not want to lose any potential customers because of the suboptimal design of some experimental pages.

Simulation (survey)-based MVLPO is built on advanced market research techniques called rule developing experimentation (RDE), a new paradigm developed in cooperation with Wharton Business School at the University of Pennsylvania (Moskowitz and Gofman, 2007).

In the *research phase*, the respondents are directed to a survey, presenting them with a set of experimentally designed combinations, namely the candidate executions for the landing page. The respondents rate each execution as a test screen on the computer, using a specific rating question and scale (*e.g.,* interest or purchase intent). At the end of the research phase, regression model(s) are created (either individually or for the total panel), relating the presence/absence of the specific features to the rating. The regression shows what features of the landing page drive the ratings. The data generate rules. These rules can be used to synthesize new pages as combinations of the top-scored elements optimized for subgroups, segments, *etc.*

This research and knowledge-building approach in most cases is much faster and easier to prepare and execute compared with the live environment optimization. It works for both high- and low-traffic Websites. Experimentation generally produces robust and rich data because of a higher level of control of the design. In contrast, there is the possibility for bias of a simulated environment as opposed to a live one and a necessity to recruit and optionally incentivize the respondents (Gofman, 2007).

The MVLPO paradigm is based on experimental design (*e.g.,* conjoint analysis, Taguchi methods, *etc.* (Green and Srinivasan, 1978)). The fundamental tenet of experimental design is to model a system by testing structured combinations of elements. The combinations reflect alternative options and are presumed to cover the "space". Of course, one cannot test tens of thousands of combinations, but one need not. The appropriate combinations can substitute for the full set of combinations that could possibly be tested.

Some vendors use a full factorial approach (*e.g.,* Google Optimizer that tests all possible combinations of elements). This approach requires very large sample sizes (typically, many thousands) to achieve statistical accuracy. Fractional designs typically used in simulation environments require testing smaller subsets of possible combinations. Some critics of fractional designs raise the question of possible interactions between the elements of the Web pages and the inability of most fractional designs to address the issue.

Advanced simulation methods based on the RDE paradigm have resolved these limitations (Moskowitz and Gofman, 2007; Gofman, 2009). RDE creates individual models for each respondent using a permuted fractional design, discovers all and any synergies and suppressions between the elements (Gofman, 2006, 2008), uncovers attitudinal segmentation and enables comparison of results across tests and over time. These features drive the MVLPO paradigm out of merely a testing system into a learning system. One learns the "rules" while doing the experiments. The first known (at the publication date) application of an experimental design to Website optimization was done in 1998 by the author of this chapter in a simulation demo project for the Lego Website (Denmark). MVLPO did not become a commercialized approach until around 2003 or 2004.

CONCLUSIONS/DISCUSSIONS

MVLPO is a powerful approach to optimize Web pages. Individual models created by RDE pave the road to real-time one-to-one marketing on Websites. The approach matches new visitors to the probable segments based on a decision tree developed during the simulation stage. This allows Website operators to *individually* optimize landing pages based on whatever information is available about the visitor [the more information that is available, the more precise may be the optimization (Moskowitz and Gofman, 2003)].

In recent years, the notion of typing the visitor has become popular as a way to individualize the landing page. The data from the RDE study―specifically the segmentation of messages― generate the necessary information with which to create a typing tool by using the method of discriminant function analysis (DFA). DFA creates a simple two- to three-question classification question, inserted at the start of a Website visit by a visitor. The pattern of the ratings to that classification identifies the segment to which the visitor belongs. Once the segment membership is known, the respondent is led to the page optimized for individuals in the specific segment. For example, in the case of food the two mind-set segments from RDE might be those interested in health *versus* those interested in convenience. A visitor filling out the typing "test" is classified immediately into one of these two segments and sent to the Web page optimized for that segment.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This chapter is partially based on Gofman, A. (2007). Consumer driven multivariate landing page optimization: Overview, issues and outlook. *The IPSI BgD Transactions on Internet Research, 3*(2): 7–9.

REFERENCES

Ash, T. (2008). Landing page optimization—The definitive guide to testing and tuning for conversions. Indianapolis, IN: Wiley and Sons Publishing, Inc.

Berland, G., Elliott, M., Morales, L., Algazy, J., Kravitz, R., Broder, M., Kanouse, D., Muñoz, J., Puyol, J., & Lara, M. (2001). Health information on the Internet: Accessibility, quality and readability in English and Spanish. JAMA, 285, 2612.

Booth, J. (2006). Can multivariate tests reduce your shopping cart abandons? Marketing Sherpa. Retrieved August 28, 2009, https://www.marketingsherpa.com/barrier.html?ident=29725

Chandler, K. & Hyatt, K. (2002). Customer-centered design: A new approach to Web usability. Saddle River, NJ: Prentice Hall PTR.

Drucker, P. (1995). Management in a time of great change. New York: Penguin Putnam.

Drucker, P. (2002). The discipline of innovation. Harvard Business Review, 80, 95–104.

Gofman, A. (2006). Emergent scenarios, synergies and suppressions uncovered within conjoint analysis. Journal of Sensory Studies, 21(4), 373–414.

Gofman, A. (2007). Improving the 'stickeness' of your website. Financial Times Press. Retrieved August 28, 2009, from http://www.ftpress.com/chapters/chapter.aspx?p=1015178

Gofman, A. (2008). Improving the "stickiness" of your website. Part 2: If they like A and B, would they like "A+B"? Wharton School Publishing. Retrieved May 15, 2009, from http://www.whartonsp.com/chapters/chapter.asp?p=1172745

Gofman, A. (2009). Experimentation-based product development in mature food categories: Advancing conjoint analysis approach. Tartu, Estonia: Tartu University Press.

Gofman, A., Bevolo, M., Moskowitz, H. (2009). Role of Corporate Leadership and Innovation Claims in Consumer Perception of Premium Products. International Journal of Innovation Management, 13(4), 655-691.

Gofman, A., Moskowitz, H., & Mets, T. (2009). Integrating science into web design: Consumer driven website optimization. The Journal of Consumer Marketing, 26(4), 286–298.

Green, P. & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. Journal of Consumer Research, 5(2), 103–123.

Janssen, K. & Dankbaar, B. (2008). Proactive involvement of consumers in innovation: Selecting appropriate techniques. International Journal of Innovation Management, 12, 511–541.

Kaushik, A. (2007). Web analytics: An hour a day. Indianapolis, IN: Wiley and Sons Publishing Inc.

Lindgaard, G., Fernandes, G. J., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & Information Technology, 25(2), 115–126.

Moskowitz, H. R. & Gofman, A. (2003). System and method for content optimization. U.S. Patent 6,662,215 B1, 2003, 16PP.

Moskowitz, H. & Gofman, A. (2007). Selling Blue Elephants: How to make great products that people want BEFORE they even know they want them. Saddle River, NJ: Wharton School Publishing.

Palmer, J. (2003). Website usability, design and performance metrics. Information Systems Research, 13, 151–167.

Schlosser, A., White, T., & Lloyd, S. (2006). Converting website visitors into buyers: How website investment increases consumer trusting beliefs and online purchase intentions. Journal of Marketing, 70, 133–148.

Temkin, B. (2007). Experience-based differentiation. Cambridge, MA: Forrester Research.

Von Hippel, E. (2005). Democratizing innovation. Cambridge, MA: MIT Press.

Von Hippel, E. & Katz, R. (2002). Shifting innovation to users *via* toolkits. Management science, 48, 821–833.

White, K. (2006). How high is your bounce rate? Retrieved August 28, 2009, from http://newsletter.blizzardinternet.com/how-high-is-your-bounce-rate/2006/02/09/

CHAPTER 17

Consumer-Driven Website Optimization

Alex Gofman[1], Howard R. Moskowitz[1,*] and Tõnis Mets[2]

[1]Moskowitz Jacobs Inc., White Plains, New York, USA and [2]University of Tartu, Tartu, Estonia

Abstract: This chapter explores one of the most advanced forms of landing page optimization: *multivariate landing page optimization*. The approach tests systematically varied prototypes of Web pages, working with consumers to identify the "rules" by which the Web pages drive specific responses. The approach, a subset of rule developing experimentation (RDE) generates indices of performance, which can translate into generalizations about consumer responses and thus "rules" about what might be done, either on an aggregated basis, on a segmented basis, or even on an individual basis. The last capability paves the road for individually optimized pages and one-on-one marketing in the near future. The approach described employs a new variation of multivariate landing page optimization based on RDE. The approach could help the marketers create better Websites that consumers like and which will help marketers to differentiate their respective Websites from their competitors.

Keywords: Conjoint analysis, experimental design, interactions, multivariate landing page optimization, regression analysis, rule developing experimentation, suppression, synergism.

INTRODUCTION

There were about two billion Internet users (as of early 2010). These users can access well over 100 million Websites. Those Websites contain tens of billions Web pages. This nets out to about five-plus Web pages for every living person! How are companies to differentiate themselves in this clutter of information?

The increase in access to information about businesses by consumers and other businesses constitutes a true double-edged sword. From one point of view, consumers gain access to thousands of new brands and millions of new items online. The traditional competition, between a few major brands distributed

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

through the local stores translates to competition in the global marketplace, empowered by the Web and turbocharged by the easily available personal computer, or cell phones with Web capabilities. The dizzying choices and global accessibility makes branding a much more difficult task.

The Internet transformed the competition into a more democratic, dynamic process and, in a sense, leveled the “battlefield” for the various players. The size and opulence of some brick-and-mortar stores might suggest their popularity, value and reach of featured brands. On the Web, it is not always the case. Many major “old” brands delayed or were late in creating a Web presence. Frequently, the resulting Websites were suboptimal, regardless of the money poured into making them “virtual stores”. Meanwhile, “new” brands jumped on the opportunity with engaging designs that gained them market share. In addition, whereas the traditional marketing approaches allocate fortunes on advertising, many brainy newcomers harness “free” word-of-mouth campaigns, taking advantages of social networking afforded by Web 2.0, *etc.* Thus, the quality and appeal of Websites no longer correlates directly with the size of a budget. The happy result for some: small and new businesses compete successfully with the giants blessed with far “deeper pockets”.

OPTIMIZING THE LANDING PAGE THROUGH EXPERIMENTAL DESIGN

In the past few years, an approach called landing page optimization (LPO) has become popular for assessing and then improving Website design. The LPO approach uses statistical design utilizing respondents who evaluate Web pages (Ash, 2008; Gofman, 2007a, 2007b, 2008). The idea behind LPO is to create several prototypes and test them with consumers. Rather than guessing, testing shows what works. Systematized testing covers more ground, shows what works and then creates rules, or at least rules of thumb.

In this chapter, we describe and explicate RDE-based multivariate landing page optimization (MVLPO) utilizing a set of experimentally designed concepts. These concepts comprise a combination of graphic and text elements, mixed and matched in order to create landing pages.

Conjoint analysis assumes that these elements constitute building blocks that appear together in a single landing page. The landing page might comprise several elements, similar to what is presented in the case study below. Test concepts represent executions of a landing page. Each test concept comprises one element from each of the silos, but other experimental landing pages might comprise fewer elements. The respondent looks at a landing page and rates the entire page on a scale such as the consumer's interest in the site, liking, or purchase intent. This single response to a landing page is an easier task than rating the different, single elements, because people are accustomed to evaluating compound vignettes (so-called gestalts).

The landing page might appear to the respondent as a set of elements combined randomly. In actuality, however, the elements are combined according to an experimental design so that:

1. each element appears equally often;
2. each element appears statistically independently of every other element;
3. a concept comprises at most one element from each of the silos;
4. in some concepts, a silo is absent (to allow for the absolute values of the utilities);
5. each element appears against many different backgrounds provided by the other elements;
6. each respondent evaluates a unique set of different combinations (each element appears several times in the combinations; the specific combination differs from person to person);
7. RDE creates an individual-level model for each respondent, showing how the individual elements drive reactions;
8. the analysis uncovers interactions between the elements as well as unique mind-set segments based on the individual patterns of responses (Moskowitz, Porretta and Silcher, 2005).

CASE STUDY

We explore RDE-based MVLPO in a study of the grocery store. As you will see, the process is very similar to the case of RDE-based package optimization (Chapter 15).

The operator of an online grocery store wanted to optimize the landing page in order to address five issues: (1) improving customers' experience; (2) enhancing the brand image of the company; (3) differentiating the brand from competition; (4) increasing the conversion rate; and (5) increasing revenue per visit.

Landing pages might have distinct layouts with a great deal of embedded information. In this simplified case, the site contains a feature picture, a banner and three different types of promotions (see Fig. **1**). These placeholders are called silos or categories (banner, feature picture). Each of the four silos on the page comprises three options, called elements. In this case, the project has four categories with three elements each. There are many more possible designs (combinations of categories and elements) readily available for different layouts. This case should be considered as a simplified sample to demonstrate the approach.

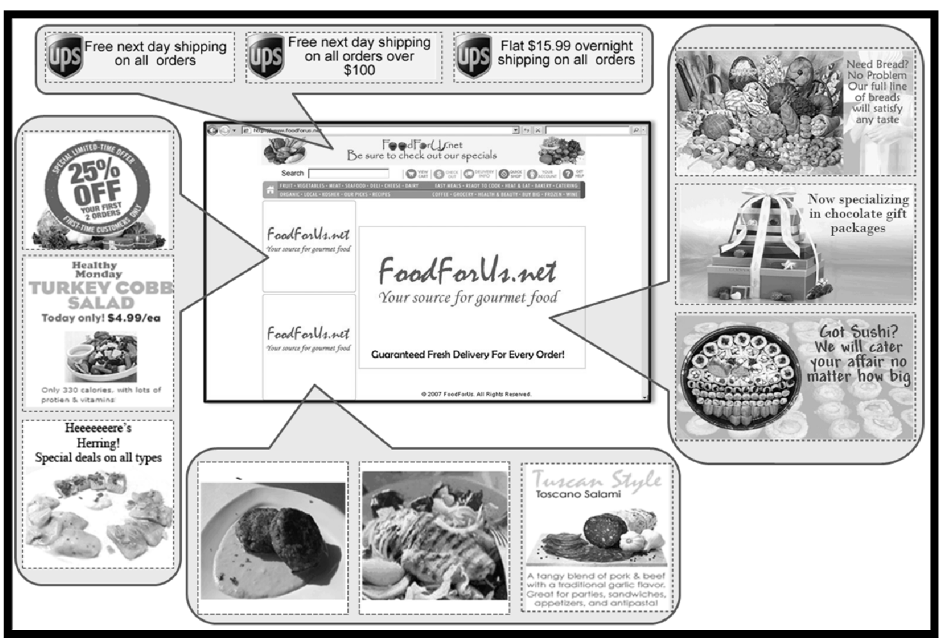


Figure 1: The template (in the middle, not to scale) and the tested elements of the Website.

The template is a schematic of the page. It places each element at a specific location. For some combinations, RDE requires that a silo be absent from the design. Although there are purists who would argue that any test Web page must contain all the components, with such a strategy one can never estimate the absolute impact of each Website element (Gofman and Moskowitz, 2010). This deliberate absence of a silo from some of the Websites enables the regression analysis to estimate the absolute values for the utility of the different design features (Moskowitz and Gofman, 2007).

For the landing page, the background of the template contains some generic text and a neutral gray background that is exposed instead of the missing options to keep the Website realistic looking.

The number of landing pages is a function of the number of silos and elements. In this particular project, RDE dynamically created a unique sequence of 27 landing page executions for each respondent. Each respondent evaluated a different set of combinations, ensuring that no specific combination of elements would influence the results.

Respondents were invited from a Web panel. A total of 1,500 invitations were e-mailed to randomly selected panel members. There were 205 responses, with 172 respondents completing the RDE survey. Each respondent evaluated 27 unique pages as dictated by the experimental design, as well as answered several demographic questions. Fig. **2** shows two sample test screens. Respondents rated each of the 27 screens on a 1–9 rating scale, answering this question: *How interested are you in purchasing GROCERIES through this Website? (1 = Not at all interested… 9 = Very interested).*

In the analysis of the landing pages, the 9-point ratings were converted to the binary scale (1–6 → 0; 7–9 → 100). The conversion is done to change intensity of interest to membership in the class "non-interested" *versus* membership in the class "interested". Following that conversion, the regression analysis showed the relationship between the presence/absence of the 12 elements and the interest in the Website. The regression, done on a person-by-person basis, showed the utility of each element and the constant.

Figure 2: Two sample screens from the interview. Each respondent has a unique set of landing page variations.

The constant, or additive constant, shows the conditional probability of a respondent being interested in patronizing the store if no elements were present. This additive constant is a purely computed parameter. The 12 different elements each generate a single utility value, showing the conditional probability of the respondent saying he or she would patronize the store if the particular visual element were present on the Website. The higher the utility value, the more likely people will buy from the store. Negative utility values mean that fewer people will buy from the store if the element were to be put into the landing page.

Table **1** presents the utility values for the different Website elements, by total panel, showing segments that emerge from the analysis, by respondent age.

1. In this case, the constant is relatively low for the total panel (+18). The constant (an estimated baseline) suggests that, in general, the respondents are not very interested in a grocery site. Only 18% would be interested in shopping at the store, without any additional information, such as the information conveyed by elements. Parenthetically, for graphics design of shopping, this additive constant of 18 is high. When we deal with packages in particular (*i.e.* small, specific items), we end up with very low additive constants, around 0. With packages it's a case of WYSIWYG ("what you see is what you get") and no more. With shopping, there is more than the Website. There is the entire experience that is expected. The higher additive constant (18) means that beyond the elements there is an expectation in the respondent's mind about the experience.

2. What could change this perception? For the total panel, only two elements substantially increase interest. One is the "Free next day shipping on all orders" and the other is the picture of the breadbasket. Other offers (middle-left panel) and food images (lower-left panel) have generally neutral utilities (from –5 to +5).

3. The data show substantial variability in utilities across the three different age groups. "The next day shipping on all orders" performs very well with younger- and middle-age consumers, increasing their purchase intent by +8 to +16 points, whereas this notion of next day shipping is actually disliked by the older audience (–8). In fact, the older consumers are not very receptive to the idea of the online grocery at all. Their additive constant is lower (+13) and most of the elements either do not change their opinion (neutral utilities) or even drive it down (negative). The exceptions are the breadbasket and chocolate pyramid pictures that cause cravings rather than making the idea of online grocery appealing.

4. Younger consumers seem to be more value-conscious as they favor "Free next day shipping on all orders" (+8) *versus* "Free next day shipping on all orders over $100" and flat shipping fees (both –2). They also like the "25% off the first two orders" offer. Nobody likes the idea of sushi delivered *via* mail.

Table 1: Performance of the Website Elements for the Total Panel, the Two Mind-Sets Segments and the Four Age Groups.

Element	**Total**	**Segments**		**Age**			
		S1: Value	**S2: Imaginary/Impulsive**	**25–40**	**41–50**	**51–60**	**>60**
Base size	172	82	90	37	63	49	22
Constant	18	20	16	21	12	25	13
Shipment options							
A_1 Free next day shipping on all orders over $100	–2	0	–4	–2	0	–3	–3
A_2 Free next day shipping on all orders	9	15	4	8	16	10	–8
A_3 Flat $15.00 overnight shipping on all orders	–3	–3	–4	–2	0	–9	–2
Promotions							
B_1 25% off your first two orders	4	5	2	8	7	–2	–2
B_2 Turkey Cobb Salad	0	2	–1	1	5	–4	–4
B_3 Herring	–1	0	–2	–3	3	–4	–2
Featured item							
C_1 100% organic range chicken	1	2	0	3	2	–4	4
C_2 Tuscan-style salami	0	3	–3	0	2	–1	–3
C_3 Quick meals	–1	2	–3	1	–1	–2	–2
Main picture							
D_1 Sushi plate	–5	–12	2	–8	–3	–8	1
D_2 Bread basket	6	–3	15	1	8	6	11
D_3 Chocolate pyramid	4	–3	11	6	6	–2	8

Note. The numbers are the conditional probabilities of a person being interested in patronizing the store if the element is shown.

Optimizing the Landing Page

One way to increase the effectiveness of the landing page is to choose the top-scoring elements from each category (Fig. **3**, left panel), using judgment to ensure that the elements "work together". The judgment is subjective, rather than defined by rules.

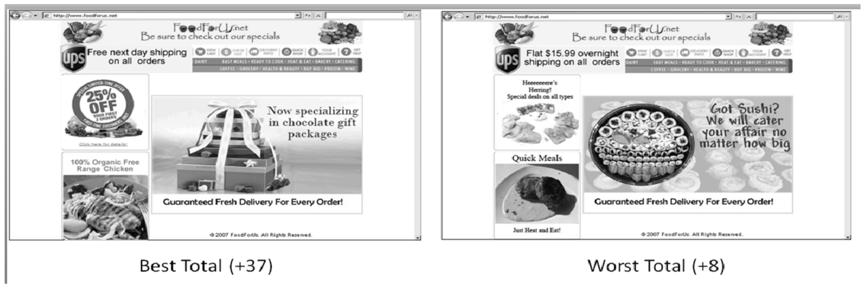


Figure 3: The strongest performing (left, +37) and weakest performing (right, +8) landing pages for the total panel, created by choosing the best *versus* the worst elements. The numbers at the bottom in parentheses are the percent top 3 box (on a 9-point scale), for respondent interest in purchasing from the Website.

The RDE utilities show the conditional probability of a person saying he or she will buy from this Website. For this study, when we select the top-performing visual elements from each category, we can increase the conditional probability of a visitor being inclined to purchase from the site from 18% (the additive constant) to 37%. The value is the sum:

$$\boldsymbol{P} = \boldsymbol{C} + \ ?\boldsymbol{u_i},$$

where P is the conditional probability of consumers being interested in buying from this site, C is the additive constant and u_i is a utility of element i.

The optimized page for these data will have the total conditional probability:

P = (Additive Constant) + Sum (Utilities) = 18 + 9 + 4 + 0 + 6 = 37.

When the "wrong" elements are chosen (*i.e.* the lowest scores in each category), the selection generates the Web page shown on the right side of Fig. **3**. The sum of utilities is far lower: (P = 18 – 3 – 1 – 1 – 5 = 8). The difference between these two choices is a possibility of losing almost 80% of potential buyers. A simple permutation of quite similar-looking elements may dramatically affect the effectiveness of the Website.

VARIABILITY ACROSS GROUPS AND MIND-SET SEGMENTS

Even more intriguing results emerge when one segments the respondents on the basis of their utilities values. People differ from each other. Males respond differently from females; high-income people perceive the pages another way *versus* middle- or low-income people, *etc.* More profound differences among people emerge when the respondents are clustered by the patterns of utilities, *i.e.* by the patterns of the elements to which they react strongly. As people's mind-sets differ, RDE reveals these different mind-sets and, thus, groups of people who differ by what drives them to like the landing page. This way of dividing people differs very much from the conventional ways that use gender, income, products purchased and the like (Moskowitz and Gofman, 2007). The RDE-based segmentation emerges after measuring the behavior toward prototype Websites and not on the basis of attitudes.

The number of segments extracted from a project using experimental design depends on at least two factors. The first factor is interpretability. The segments must be interpretable, which comes most easily when the study involves different kinds of elements, specifically dealing with topic areas. It becomes easy to identify a common thread when segments divide by the topics that interest them. Furthermore, there is no reason to extract more than the minimum number of segments. In the interest of simplicity and parsimony, the fewer the segments extracted, the stronger the data and the more cogent the results. Thus, when we find that two mind-set segments "tell the story coherently," then there is no reason to create more segments. In contrast, there are occasions when three or more segments are needed because the two-segment solution is simply not sufficiently coherent for one or even both segments. In the extreme case, one might even create an optimized page for each individual, a one-person segment, facilitating real one-to-one marketing (Moskowitz and Gofman, 2003).

The second factor is sample size, which should be sufficiently large so that a split of the same into different groups will still provide data from each group that is robust. At a minimum, about 100 respondents should participate if one expects to segment data. The sample size should be big enough to be split into meaningful segments with a feasible number of intended customized pages, *etc.*

In the case study, two approximately equally sized mind-set segments emerge. We called Segment 1 the "value oriented" group and Segment 2 is "imaginary/impulsive" group. The segment names come from the elements that score best. Looking at Figs. **4** and **5**, we get a sense of what silos perform well.

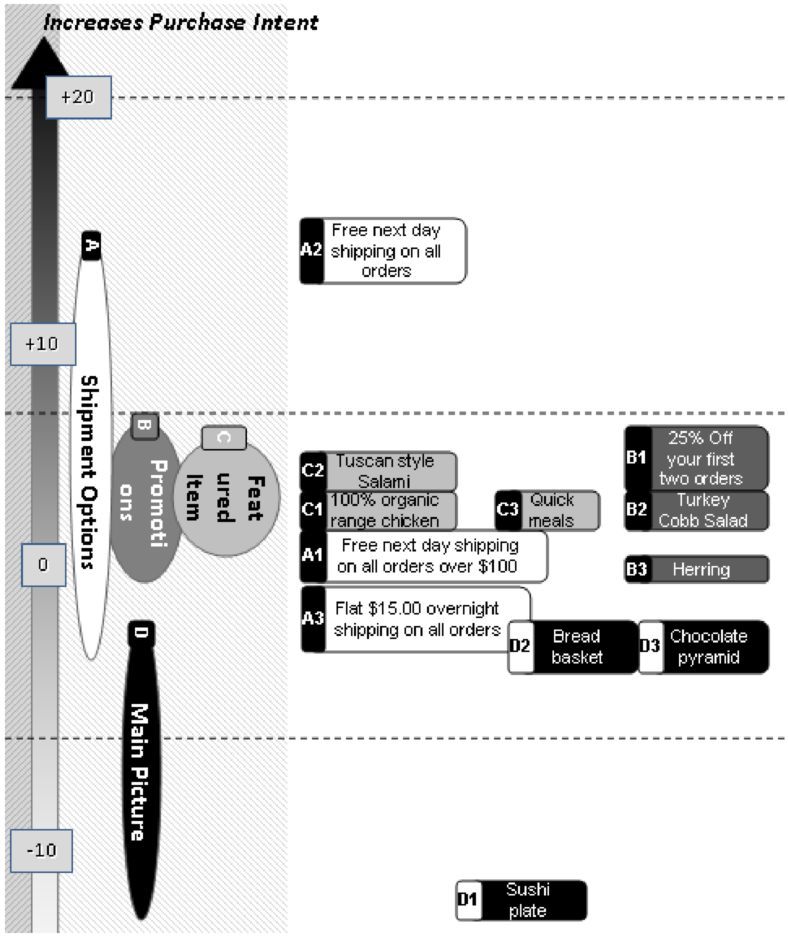


Figure 4: Performance of elements for the Segment 1 (Value Oriented). The categories appear as ovals. Their vertical size represents the spread of impacts of the elements that belong to the category. The elements (rectangles) are located on the vertical scale of impacts (from negative at the bottom to positive on top).

Granularity counts in segmentation, rather than generalities. A single silo comprises several elements, some of which may do well, whereas others may do poorly. In segmentation, the real meaning lies in the individual elements rather than in some more general, overarching rule.

1. Segment 1, the value oriented, shows a higher constant. This suggests that this segment is a bit more positive to the general idea of the online grocery. Respondents belonging to Segment 1 love free next day shipping (+15) and are positive to the offer of 25% off (+5). The

remaining elements are close to 0, so they feel it is irrelevant, with one exception. The picture of sushi is a strong negative (–12). Segment 1 apparently believes that such perishable food items as sushi should not be even offered by an online shop.

2. Segment 2, the imaginary/impulsive, is generally neutral to everything except for the central mouth-watering images of a breadbasket (+15) and chocolate pyramid (+11). Segment 2 is far less receptive to value offers, suggesting that they respond to the items, not to cost savings.

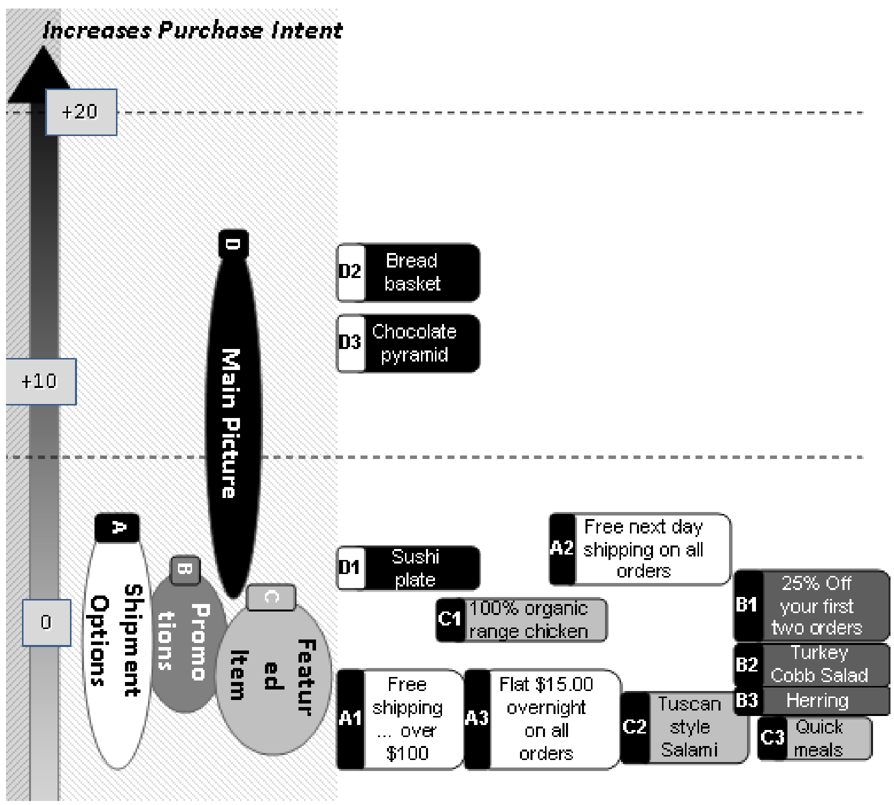


Figure 5: Performance of elements for the Segment 2 (Imaginary/Impulsive).

Optimizing the landing page is considerably easier for these segments because they exhibit coherent, more single-focused responses. The segments were developed to be homogeneous in terms of the patterns of their responses. We see the optimized landing pages for the two segments in Fig. **6**.

Figure 6: Optimized landing pages for the Segment 1 (left) and Segment 2 (right).

Improving Landing Pages by Using Positive Interactions Between Its Elements

In a number of occasions, some latent but often strong interactions exist between the tested elements on the page. Using a highly trained expert's opinion to guess these interactions—a traditional approach—is not a very viable option in the fast-moving world of Website design. RDE overcomes the limitations of the old methods by automatically testing all pair-wise combinations of the elements of the page according to a built-in, permuted experimental design, which allows the discovery of significant interactions, both of positive and negative natures, respectively.

The process to discover the interactions (synergisms, suppressions) follows the four steps (Gofman, 2006, 2008, 2009; Moskowitz *et al.* 2005).

1. *Data preparation.* Create a matrix of all the raw data, comprising rows of 12 columns each (one per each element). With 27 landing pages tested by each participant and with 172 participants, the data matrix comprises 4,644 rows.

2. *The purchase intent measure.* Create the 13th data column that takes on value 100 when the respondent rated the landing page 7–9 to denote positive interest. The 13th column takes on the value 0 for the rating is 1–6, to denote lack of such intent. It is done for easier interpretation of what the rating value means (intensity of feeling *versus* membership in a specific group, *i.e.* not buy *vs.* buy).

3. *Create all pairs of elements from each of the two silos*. There are four silos, A–D, so there are 6 pairs of silos [(4 × 3)/2 = 6]. Each pair of silos generates 9 pairs of elements (*e.g.,* A1 … A3 crossed with B1 … B3 generates 9 combinations). Therefore, there are 6 pairs of silos × 9 combinations for each pair, or 54 pairs of elements.

4. *Identify the pair-wise interactions co-vary with interest*. First create a model relating the presence/absence of the 12 single elements to the dependent variable. *Force those 12 elements into the equation*. Then for the remaining 54 predictors corresponding to the interactions, look at which elements explain additional variability.

Before creating a new combination of ideas by splicing together components, it is important to determine whether the combinations "work" together or whether they do not. Some combinations make intuitive sense, while some do not. These combinations may be identified ahead of time and specified as pair-wise restrictions. However, there are many combinations that simply cannot "work" together, even though there is no *a priori* reason to assume that they would fail to work. It may be that the combinations are counterintuitive or clash with each other.

Not every case produces meaningful interactions. On many occasions, interactions are not very strong and could be ignored (considered not significant). When the utility (conditional probability of customers being interesting in buying from this site) of the combination falls inside the empirical "neutral" range of (±5), it could be discarded.

It also should be noted that the effect of the interactions changes the regression model and somewhat affects the rest of the utilities. In a model without interactions, the values of hidden synergies and suppressions are disseminated among the individual elements and cannot be uncovered. In a more comprehensive regression model that includes interactions, the values are extracted and assigned to the cross-terms (pairs of elements).

Table **2** shows the utilities of the individual elements of the landing page and finishes at the bottom with a discovered meaningful interaction (the last row).

Table 2: RDE Model with Interactions.

Element	**Utility**
Constant	15
A_1	–1
A_2	12
A_3	–2
B_1	3
B_2	1
B_3	–2
C_1	2
C_2	1
C_3	2
D_1	–4
D_2	7
D_3	5
A_2C_3	–6

Note. The utilities are slightly different than in the standard model.

The only moderately meaningful interaction between the elements A_2 and C_3 is negative (–6), meaning that the combined effect of these two elements will generate suppression. The effect of combining these two elements will not be:

(Combined effect of A_2 and C_3 w/o interactions) = $A_2 + C_3 = 12 + 2 = 14$,

but rather:

(Combined effect of A_2 and C_3 w/interactions) = A2 + C3 + A2 * C3 = 12 + 2 - 6 = 8.

This clearly should be taken into account during the optimization process.

This specific case did not produce very strong interaction values and the analysis has discovered only one such interaction, but it does demonstrate the approach. In some other cases, an interaction along could add 20 or more points to the score (or subtract from it). By looking into this wealth of data created automatically by RDE tools, we can greatly improve customers' experience and increase their purchase intent.

Individual models afforded by RDE also pave the road to real-time one-to-one marketing on Websites by matching new visitors to the probable segments. This would allow Website operators to individually optimize landing pages based on whatever information is available about the visitor [the more information that is available, the more precise may be the optimization (Moskowitz and Gofman, 2003)].

CONCLUSIONS

Consumers might not know what they want deep inside, but they will react when given alternative options. This notion recently led to an explosion of experimentation as a method for optimizing Web design. To find a winner, one should experiment and test multiple, systematically varied prototypes in order to identify patterns of stimuli that drive responses. A simple permutation of similarly looking elements may produce a noticeable and important impact on the purchase intent of consumers and perhaps on the subsequent conversion rate.

Conventional methods using focus groups and simple concept tests often turn into simple "beauty contests", which produce a simple answer but nothing more. These research methods generally cannot create rules of the consumer mind. Furthermore, because they are limited to a few executions, they generally are an inefficient expenditure of research funds. They can only find winners among the stimuli tested. They cannot really educate for further efforts, except by happenstance. In contrast, RDE-based MVLPO systematically tests hundreds or even thousands of landing pages, develops rules, identifies segments, detects interactions and generates winning propositions. The goal of RDE-based MVLPO is to design/test/learn/create rules/optimize. Essentially, the approach constitutes a virtuous circle, with ever-increasing understanding.

With the introduction of new tools, MVLPO has made the field more democratic, available to virtually any Website operator. For professional designers, RDE tools create an opportunity to improve their designs even further. The research efforts generate solid consumer data, which reveal the anticipated effects of the design. The output of RDE should be taken as the input for the artistry of the designer, who adds the individual expression. The combination of designer and RDE-based consumer insights might well create the best of all worlds—art, design and science combined to improve Websites and thus the experience of the Website's visitors and ultimately the bottom line of the business.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

[1]This chapter is partially based on Gofman, A., Moskowitz, H. and Mets, T. (2009). Integrating science into web design: Consumer driven website optimization. *The Journal of Consumer Marketing, 26*(4), 286–298.

DISCLOSURE

Part of the information included in this chapter has been previously published under the title Integrating Science into Web Design: Consumer Driven Website Optimization. The Journal of Consumer Marketing, (2009), 26(4), 286-298.

REFERENCES

Ash, T. (2008). Landing page optimization: The definitive guide to testing and tuning for conversions. New York: Wiley and Sons, Inc.

Gofman, A. (2006). Emergent scenarios, synergies and suppressions uncovered within conjoint analysis. Journal of Sensory Studies, 21(4), 373–414.

Gofman, A. (2007a). Consumer driven multivariate landing page optimization: Overview, issues and outlook. The IPSI BgD Transactions on Internet Research, 3(2), 7–9.

Gofman, A. 2007b. Improving the 'Stickeness' of Your Website. Wharton School Publishing, http://www.whartonsp.com/articles/article.asp?p=1015178, September 21, 2007 (accessed on May 15, 2010).

Gofman, A. 2008. Improving the "Stickiness" of Your Website. Part 2: If They Like A and B, Would They Like "A+B"? Wharton School Publishing, http://www.whartonsp.com/articles/article.asp?p=1172745, February 22, 2008 (accessed May 15, 2010).

Gofman A. (2009). Experimentation-based product development in mature food categories: Advancing conjoint analysis approach. Tartu, Estonia: Tartu University Press.

Gofman, A., Moskowitz, H. (2010). Application of Isomorphic Permuted Experimental Designs in Conjoint Analysis. Journal of Sensory Studies, 25(1), 127-145.

Moskowitz, H. R. & Gofman, A. (2003). System and method for content optimization; U.S. Patent 6,662,215 B1, 2003, 16PP.

Moskowitz, H. & Gofman, A. (2007). Selling Blue Elephants: How to make great products that people want before they even know they want them. Upper Saddle River, NJ: Wharton School Publishing.

Moskowitz, H. R., Porretta, S., & Silcher, M. (2005). Food product design and development. Ames, IA: Blackwell Publishing.

CHAPTER 18

Consumer-Driven Website Customization: The Need to Manage Costs and Benefits

Eelko K.R.E. Huizingh[1,*] and Marije L. Teerling[2]

[1]Department of Innovation Management & Strategy, Faculty of Economics and Business, University of Groningen, The Netherlands and [2]Telematica Instituut, Enschede, The Netherlands

Abstract: Website customization is an innovative Internet concept to optimize consumer experiences. By pursuing three objectives, this study aims to increase our understanding of how and when site customization can be applied. First, we define the concept of Website customization as an extension of mass customization. Second, we identify the value drivers of Website customization from both a customer's and a supplier's point of view. Finally, we investigate how two companies, which operate in different environments, deal with these value drivers. The two case studies, concerning Amazon and Dell, focus on companies that are widely recognized as being successful in this area, but operate in significantly different environments. The case analyses show many similarities in the customization strategies of both companies, *e.g.,* delivering substantial added customer value, automating the customization process, offering a stepwise process and the crucial role of trust. However, there are also distinct differences regarding the initial customer's investment, the elements of added value and the knowledge on which customization is based.

Keywords: Case study, e-commerce, marketing interface, mass customization, website.

INTRODUCTION

The Internet can have a significant impact on the strategic positioning of organizations (Porter, 2001). As such, many companies are using the Internet to implement innovative strategies. When searching for opportunities where they can apply the Internet, companies can focus on current customers to deepen customer relationships, on new customers to develop new markets, on their strategic positioning to strengthen their business network, or on creating new products

***Address correspondence to Eelko K.R.E. Huizingh:** Department of Innovation Management & Strategy, Faculty of Economics and Business, University of Groningen, The Netherlands; Tel: +31 (0)50 363 91 11; Fax: + 31 (0)50 363 53 80; Email: k.r.e.huizingh@rug.nl

Alex Gofman and Howard R. Moskowitz (Eds)

and/or new value-adding services (Huizingh, 2002). In this chapter, we focus on the potential of the Internet to enhance mass customization strategies.

Kiang, Raghu and Shang (2000) considered the potential of customization as the most important factor when determining the suitability of the Internet for tangible products. In their field study of Internet-based businesses, Grover and Saeed (2004) reported that the cluster with the most successful companies contained many businesses well known for their high level of customized (personalized) services. Customization enabled these companies to achieve both efficiency and effectiveness in leveraging their customer base and generated higher margins. Not surprisingly, we observe a fast-growing interest in site customization by practitioners. A study by Jupiter Research found that 38% of the surveyed US companies already had invested in customization, whereas another 35% have planned personalization initiatives in the next 12 months (Surmacz, 2003). Liao, To and Shih (2006) found that more than 70% of the US Fortune 1000 companies are using cookies to collect user profiles for customization. Customization is widely recognized as a promising concept to innovative marketing strategies, as witnessed by publications of leading practitioners (Kasanoff, 2002; Nilson, 2002).

Despite the interest in site customization, Huang and Lin (2005) observed that many personalization-related concepts are still forming. Companies are still in the phase of experimenting with these features. Franke, Keinz and Steger (2009) stress the importance of the ability to obtain precise information on what customers actually want. Teo (2005) reported considerable interest in customization features among Singapore firms but also found a significant difference between its mean extent of use and its perceived effectiveness. Franke and Piller (2004) discuss the recent shutdown of two well-known mass customization Websites, Mattel's *My-Design Barbie* and Levi's *Original Spin* site; and Taylor, Terwiesch and Ulrich (2005) discuss methods to overcome the limitations of such sites. Finally, a market study of Jupiter Research questions the effectiveness of site customization by stressing its high costs (Surmacz, 2003). In summary, Website customization seems to be a mixed blessing that could be compared with IT innovations. Apparently, customer involvement is an important factor (Jayawardhena and Wright, 2009).

The foregoing conclusion about the nature of customization as a “mixed blessing” underscores the need to increase our understanding of the concept of site customization and especially how and when it can be applied in practice. Under which circumstances may we expect successful implementation? How can companies deal with both the site customization cost and benefits? When is it likely that customers will use the customization features offered in a site?

Our study aims to address these issues and contributes to the emerging stream of site customization research in three ways:

1. We define the concept of Website customization as an extension of the mass customization concept.

2. By combining a rational actor theory, *e.g.,* transaction cost economics (Williamson, 1975), with a relational approach, we identify the value drivers of Website customization from both a customer’s and a supplier’s point of view. It is essential to uncover both parties’ value drivers, as only in case of a real win–win situation will site customization flourish.

3. In the empirical part of our study we apply the value drivers model to two case studies (Amazon and Dell) to investigate how both companies have implemented the value drivers in their customized sites. We selected these two companies because they are highly regarded for their customized Websites, yet they operate in fairly different environments (the consumer market and the business market), which makes it interesting to compare their approaches.

FROM MASS CUSTOMIZATION TO SITE CUSTOMIZATION

Mass customization arrived on the scene of management theory and practice in the late 1980s (Da Silveira, Borenstein and Fogliatto, 2001). A visionary definition of mass customization describes it as a concept to provide customers with anything, anytime, anywhere in any way they want it (Hart, 1995). Mass customization is characterized by the involvement of the customer in the design, production, or delivery process before the actual transaction takes place (Kamali and Loker, 2003). With the advent of the Internet, even more applications of customization are expected (Swaminathan, 2001), as any Website relaxes the time and place restrictions (“anytime and anywhere”). Website customization offers the opportunity to tailor the contents of a Website to the specific needs of a customer (“any way they want it”).

One of the strengths of the Internet is that it is a two-way medium. Mass customization requires two-way media, as customers have to be able to express their individual needs and wants to the supplier (Bardakci and Whitelock, 2003). The importance of two-way media is often stressed in marketing literature. New communication technologies enable companies to prepare, on a mass basis, individually designed communications to meet each customer’s requirements (Kotler *et al.* 2003).

The rapid development of the Internet as a tool for commercial purposes, combined with the developments in the area of mass customization, has led to the emergence of Website customization. Researchers have recognized the innovative potential of the Internet to tailor products, services and the transactional environment to individual customers (Angehrn and Meyer, 1997; Rowley, 2002; Srinivasan anderson and Ponnavolu, 2002; Wind and Rangaswamy, 2001). In line with mass customization definitions, Ansari and Mela (2003) further specify site customization as the extent to which suppliers either customize the Website to appeal to users or enable the users themselves to customize the content. Similarly, Teerling and Huizingh (2006) define Website customization as the extent to which a Website contains pages that are tailored to or by individual customers.

COSTS AND BENEFITS OF SITE CUSTOMIZATION

The more critical discussions of Website customization show that it is a mixed blessing. As with any IT innovation, it comes with both costs and benefits. In this section, we identify the various cost and benefit elements of site customization. We follow the accepted approach in information systems (IS) research on the role of IT in mediating customer–supplier relationships by combining rational actor theories, *e.g.,* transaction cost economics (Williamson, 1975), with relational approaches (Schultze and Orlikowski, 2004). This enables us to construct a more complete image of the factors that are relevant, as rational, transaction-oriented theories tend to underestimate the importance of softer, more long-term-oriented factors. We identify value drivers from both a customer's point of view and a supplier's point of view as both parties obviously have different concerns.

The Customer's Point of View

In most cases, customers have the option to use either the standardized pages within a site, or to use the customized pages. If customers of a book site decide not to provide the e-tailer with their preferences, they still have access to all the books in the site. This implies that customers have to make the decision whether site customization is worth the additional effort. They will make the trade-off between the costs and benefits of site customization (see Table **1**).

Table 1: The Benefits and Costs of Site Customization for Customers.

Benefits	Costs
Better information and time savings	Time (investment)
Useful recommendations	Cognitive efforts
Increased sense of control	Privacy violation
Sense of esteem	Exploitation by supplier

Benefits of Site Customization

1. *Better information and time savings*. Site customization enables customers to become better informed, in terms of preferences, interests, or assets. Customers get more complete and relevant information that can be accessed faster (Hoffman, Novak and Chatterjee, 1995). Instead of being presented with all available products, a customer is confronted with a list of the most relevant products only (Vrechopoulos *et al.* 2003). This narrowed list of choices improves customer decision making and creates a more efficient buying process through faster provision of relevant information (Burke, 2002; Reibstein, 2002).

2. *Useful recommendations*. The more complex a product and the less product knowledge a customer has, the more difficult the purchase decision becomes. A customized site may act as a smart salesperson offering a customer relevant recommendations. Customization is

particularly useful when customers are not able to fully or accurately self-explicate their preferences (Ansari and Mela, 2003). Customers can also gain from recommendations in situations where a large number of options are available. A large number of options increase the perceived complexity, which may lead to suboptimal purchase decisions (Godek *et al.* 2002; Reibstein, 2002). Kamali and Loker (2003) found that a customer is just as satisfied with 50 options as with 37,500. Therefore, the purchase decision can be improved by applying customization to recommend a specific product or recommend a relevant subset of all available options.

3. *Increased sense of control*. Site customization may increase the sense of control that customers experience while using the online environment (Dann and Dann, 2004). Site customization enables customers to control, at least to some extent, the contents of a site. Instead of a site showing the market prices of all kinds of stocks, customers can make the site show the value of only the stocks they own. In an empirical study, Nunes and Kambil (2001) found that customers value the opportunity of being in control.

4. *Sense of esteem*. Site customization provides customers with a sense of esteem from the organization. The site greets the customer with his/her own name and acts as an informed and understanding assistant by selecting items that are relevant to this particular customer. In so doing, the organization recognizes the customer as an individual with unique needs and wants. The site acknowledges him/her as an important customer, which may result in an enhancement of customer commitment and trust.

Costs of Site Customization

Zeithaml (1988) distinguishes between two types of customer sacrifices, monetary costs, the price of a product and non-monetary costs, *e.g.,* time, search and psychological costs. Site customization usually only involves non-monetary costs. The costs are affected by the following four issues:

1. *Time*. In order to use site customization features, customers often fill in a registration form and answer a number of questions with regard to their preferences, interests and/or assets.

2. *Cognitive efforts*. Using customization features requires that customers be willing and able to learn how the supplier has implemented these features in the site. Even when a customer uses customization features in various sites, the learning effects across sites are limited. Each Website implements the customization features in a different way. Some customers consider learning to deal with new technologies to be confusing process, hindering adoption (Mick and Fournier, 1998).

3. *Privacy violation*. The use of site customization requires customers to provide the supplier with personal information. Depending upon the degree of sensitivity of this information, the site may face privacy concerns. Several studies found that customers have reservations concerning the collection and use of personal information in order to tailor marketing programs (Burke, 2002; Schafer, Konstan and Riedl, 2001; Schoenbachler and Gordon, 2002; Wolfinbarger and Gilly, 2003), as it is possible that suppliers will use the personal information in an unintended way. Nunes and Kambil (2001) consider privacy concerns as the leading reason for consumers to withhold information from sites, making trust an essential element of online transactions (Bunduchi, 2005).

4. *Exploitation by supplier*. When a supplier knows more about a customer, the supplier's control over the customer increases. For instance, when a supplier discovers that a customer is not price sensitive, the supplier may decide not to offer any discounts any more (Murthi and Sarkar, 2003). Site customization then turns into a win–lose situation, instead of a win–win situation. Customers have to trust the supplier in order to accept this risk.

The Supplier's Point of View

Suppliers face multiple (possible) benefits and costs of site customization (see Table **2**). We first discuss the various five categories of benefits and then the four costs of site customization.

Table 2: The Benefits and Costs of Site Customization for Suppliers.

Benefits	**Costs**
Site-related benefits	Initial investment
Transaction-related benefits	Knowledge
Information-related benefits	Maintenance costs
Relationships-related benefits	Failure to meet increased customer expectations
Increased switching costs	

Benefits of Site Customization

1. *Site-related benefits*. Increased customer site satisfaction may lead to more repeat visits and increased stickiness, *e.g.,* longer visits or more page requests. Several authors assume that the stickiness of the Website can be improved by using customized Web pages (Johnson, Bellman and Lohse, 2002; Thompson, 1999). When the site receives revenues from advertising, then the site-related benefits have direct monetary consequences.

2. *Transaction-related benefits*. Site customization may result in more purchases and more frequent purchases. When the customization features streamline the ordering process, site customization may be instrumental in solving the well-known problem of the "abandoned shopping carts" online (Bizrate, 2000; eMarketer, 2004). Furthermore, improved knowledge about customers' preferences may allow the organization to more effectively cross sell or up sell products. Customized e-mailing may increase efficiency by reducing transaction costs (Grover and Saeed, 2004).

3. *Information-related benefits*. Site customization not only requires intimate customer knowledge (Vrechopoulos *et al.* 2003) but can also be considered as a source of customer information (Hanson, 2000; Thompson, 1999). One of the limitations of database marketing is that suppliers know a lot about what, when and how many customers buy, but not much about why they buy. Site customization enables suppliers to gain a better understanding of their customers. This

knowledge can be leveraged to develop more effective marketing programs with better-targeted offers.

4. *Relationship-related benefits*. Site customization leads to a more individualized online experience, which in turn may improve customer satisfaction and loyalty (Bolton, 1998; Grover and Saeed, 2004; Peppers and Rogers, 1999). Some authors consider this potential benefit to be one of the greatest assets of site customization (Reichheld and Schefter, 2000). Suppliers can try to exploit online relationships, for example by using customer profiles to select customers for participation in the development of a new product that complements their current purchases.

5. *Increased switching costs*. Suppliers can use site customization to try to lock in customers by increasing their switching costs. A supplier may initially offer site customization with a low threshold (*e.g.,* by asking for only a few personal characteristics and offering an automatic log-in function), while gradually, over time, the customers can be asked for more personal details. When the site experience is improved accordingly, the supplier increases the switching costs step by step. In addition, customers will realize that in order to get a similar experience on a competitor's site they have to provide the same amount of information again. Switching costs also increase when customers get used to the tools offered on a site, or when a supplier offers, for example, membership credits to reward customers for certain actions.

Costs of Site Customization

1. *Initial investment*. When site customization features are added to the site, the supplier not only has to develop tailor-made Web pages, but it also needs an authentication process, databases to store the customer information and systems to link the customer information in a meaningful way to their offerings. Site customization involves many additional operations for each site visit, likely leading to requests for more advanced hardware to ensure acceptable loading times. One

study found that operating a personalized Website can cost upwards of four times that of operating a comparable dynamic site (Surmacz, 2003).

2. *Knowledge*. Developing effective site customization features requires knowledge from at least three different sources:

    a. Customer knowledge is needed to understand why and how customers intend to visit the site.

    b. Statistical knowledge to infer relationships between customer characteristics and customer behavior (both site browsing and purchasing).

    c. Technological knowledge to develop the proposed site customization features and to effectively link them to the existing IT systems.

3. *Maintenance costs*. Like all systems, site customization features have to be maintained. Maintenance includes updating dynamic personal information and developing offerings that are tailored to particular customer profiles. Suppliers have to realize that introducing site customization features is not a project but a process, probably one that never ends. A Jupiter Research report describes a health-care site that spent 5 months weaving scenario management through its entire site, only to realize that maintaining the staff required to manage campaigns, scenarios and rules was prohibitively expensive (Festa, 2003).

4. *Failure to meet increased customer expectations*. When a supplier asks for personal information, customers will expect to receive more relevant site content in return. In a class one of the authors was teaching, students analyzed the customization features of an e-tailer. One student provided the e-tailer with the information that she was a lady of over 80 years. During one of her next visits, she was provided with “a personalized offer” for a skiing outfit. If a supplier is not able

or does not know how to meet the increased customer expectations, the supplier should refrain from offering customization features to avoid these potential costs.

SITE CUSTOMIZATION CASES: AMAZON AND DELL

We selected two companies, Amazon and Dell. Both companies are widely recognized as being successful in the area of site customization, but operate in significantly different environments. Amazon is one of the leaders of customization in the consumer market, whereas Dell's Premier Service represents a major example of site customization in business markets. As both companies are assumed to be well-known, we need not provide further introduction. This section provides a description of both cases, mainly based on information from both sites (http://www.amazon.com and http://www.dell.com); in the next section, we analyze how Amazon and Dell deal with the different value drivers.

Site Customization at Amazon

Customization is at the heart of Amazon's strategy, as reflected by an often cited quote of CEO Jeff Bezos: "Our vision is that if we have 20 million customers, then we should have 20 million stores". Collaborative filtering smartly combines browsing data from one customer to data from other customers to offer functions such as "Customers who bought this item also bought…" and "Customers with similar searches purchased…"

Amazon collects data about items customers search, add to their wish list and buy. At many instances the site uses these data, *e.g.,* the home page shows books the customer browsed during a previous visit and announces the availability of new recommendations. The customer's personal page includes many links to customized pages such as an overview of recent purchases, "Today's Recommendations for You," "Your Recently Viewed Items," "Coming Soon for You," and "New for You".

In addition to observing customers, Amazon frequently asks for customer feedback, *e.g.,* a rating of a recently purchased item or whether a help page was useful. Amazon also tries to start dialogues with customers. Customers can ask why

Amazon recommended a particular item, Amazon answers by displaying the purchases that led to this recommendation. Customers are then asked to rate these items, but can also indicate that Amazon should not use an item for making recommendations (*e.g.,* because the purchase was a gift). Customers can rate recommendations or indicate that they already own a recommended product. The information collected in the dialogues is used to further improve the recommendations.

Another way to increase customer involvement and collect customer information offers personalized weblogs (called "plogs"). The weblogs contain messages concerning recently released items, changes in orders, or postings from Amazon editors or authors that are related to the customer's purchases or searches. Just like with regular weblogs, Amazon customers can comment on the messages.

Amazon also enables customers to interact with other customers. Customers can request the wish list of other customers. They can also create a wedding registry or baby registry. They become part of the Amazon community by creating a profile, including personal information such as a birth date, interests and a list of favorite items they own. For some personal data, customers can limit access (to the customer only, friends only, or everyone).

Site Customization at Dell

The Dell Premier Service enables customers to manage all phases of computer ownership. Dell's customized secure site offers purchasing, order status reporting, management reporting, service and support. The core of the service, initially called Dell Premier Pages, is a customized online computer store. The customer's IT managers can create and maintain standards by pre-designing configurations and selecting approved supplies and accessories. The customer's employees can directly order these products. When entering the online catalog, they are only offered the pre-approved options with the pre-negotiated prices. As different departments may have different IT needs, it is also possible to define distinct standards for specific user groups.

Dell Premier streamlines the customer's purchase process. End users can configure and price their own systems and save them as e-quotes. Authorized buyers are then notified by e-mail and asked to give approval and submit the order to Dell. As the system is aware of spending limits, it first checks whether budget is available. The order tracing and tracking function produces e-mail order confirmation, real-time visibility into the order status and shipment notification.

An important aspect of purchasing management is the definition of various roles to individuals in the organization. Dell Premier distinguishes between six levels of accountability, which determine what an individual user can do and see. The roles range from shoppers, who can view all the content for their access group and save e-quotes but not buy anything, to buyers who can also place orders.

Dell Premier keeps track of all PCs purchased by a customer and the parts added to these machines. In this sense, the service acts as an asset management system. Dell also uses this information proactively, *e.g.,* by informing customers about which particular machines need to download a software fix.

The system offers various management reports for purchase and asset management. The reports offer insight on current orders, shipping lead times, past invoices and how much the company is spending on various Dell products. The system contains both standard reports and features to build custom reports that can be saved for ongoing use.

Additional functions include ImageWatch, with information about future technology changes, technical support, warranty expiration and upgrade notifications and the Dell Knowledge Base that gives instant answers to technology questions from a database of problems and solutions. Dell even offers to fully integrate the Premier site with the customer's existing procurement system.

The three major groups in the customer's organization benefiting from Dell Premier are end users, the IT department and the purchasing department. End users do not have to worry about whether the products they intend to order are approved and gain from faster ordering. IT departments use Dell Premier as an efficient and effective tool to manage IT standards. Purchasing departments save on operational issues (*e.g.,* answering questions on issues that have already been negotiated) and use the service as a tool to enforce controls related to the what, who and how much of purchasing.

VALUE DRIVER ANALYSES OF BOTH CASES

Based on the framework developed earlier in this chapter, we now consider how both companies deal with the value drivers, from the respective viewpoints of the customer and the supplier. These analyses show how both companies maximize the benefits and minimize the costs of Website customization.

The Customer's Point of View

Benefits of Site Customization for Amazon and Dell Customers

1. *Better information and time savings*. At Amazon, better information implies better recommendations. Customers save time by not having to browse through a vast amount of available items and by using Amazon's patented one-click ordering procedure. At Dell major savings come from customers seeing only products approved by their IT department, simplified and shorter purchasing process cycles and the reduction of errors in this process.

2. *Useful recommendations*. Amazon is highly regarded for the quality of its patented recommendation algorithms by which it is able to make useful recommendations even for low-volume products, the so-called long-tail items (Anderson, 2006). The Dell recommendations are valuable as customers know that these items are approved by their IT department. Dell Premier is less concerned with helping to find the best configuration given an end user's needs and wants.

3. *Increased sense of control*. Amazon searches for the delicate balance between providing effective recommendations and increasing the customer's sense of control. Amazon does so by offering customers the opportunity to turn off some recommendation features. For example, customers can temporarily turn off their browsing history, delete all items from their browsing history, tell Amazon to not use a purchase for making recommendations, or remove a message from their personal weblog. Dell customers are empowered to maintain the catalog of approved items, to determine the purchasing roles of employees, to set budgets and to customize management reports.

4. *Sense of esteem*. Amazon greets customers by name on its home page, offers relevant recommendations and keeps track of confidential personal data (such as multiple credit cards or delivery addresses) while still being sensitive to user requests (*e.g.,* to not use a particular purchase for making recommendations). Dell Premier comes in a customized look and feel including the customer's logo. Whereas Amazon offers its customization features to any customer, Dell focuses on a select group of customers and only offers its customized service to medium and large businesses and institutions.

Costs of Site Customization for Amazon and Dell Customers

1. *Time*. When setting up the system, Amazon's customers do not have to spend time; the site learns preferences by recording browsing and purchasing behavior. Even logging in takes place automatically, as Amazon uses cookies to recognize customers. Customers do have to login when purchasing, however. In several instances, Amazon asks customers for additional data, *e.g.,* ratings of recommendations and purchased items, products they own, or to mark gifts. In contrast, Dell's customers have to spend considerable time setting up the system. They have to provide information such as approved items, details of their procurement process and authorization levels of employees. However, in both cases customers can start at a low level, Amazon customers do not have to respond when Amazon asks for additional information and Dell customers can use Dell Premier as no more than a customized catalog.

2. *Cognitive efforts*. Amazon tries to make the customization barriers as low as possible. Except for having to remember their password, no cognitive efforts are required. In the case of Dell, major cognitive efforts are required due to the high level of formalization required. Before using the service, a corporate customer must have approved configurations, but there might also be a gray area with configurations normally not approved but that sometimes are, providing the right person is approached by the right person with the right arguments.

When Dell Premier lists only the approved configurations, this gray area disappears. To help customers overcome such barriers to setup, Dell representatives assist in this task. That effort is the reason why Dell only offers the Dell Premier Service to its larger customers.

3. *Privacy violation*. Amazon collects and stores any (personal) information customers enter or provide in any other way. As Amazon uses cookies to identify individuals, anyone having access to a customer's PC has access to this person's recommendations. However, a customer can decline to provide Amazon with any additional data it asks for and indicate he does not want to receive e-mail offers from Amazon or other related businesses. Users of Dell Premier have to reveal many details of their purchasing process, including the authority levels of various employees and their spending budgets. As they are organizations, privacy is not so much an issue, but due to the confidentiality of the information, security is.

4. *Exploitation by supplier*. In both cases, trust plays a major role. Amazon customers have to trust that the company is not abusing their personal information, which Amazon (in theory) can easily do. For example, information about an individual's price sensitivity could be used to offer personalized prices. The author once experienced that when accessing the Amazon.uk site for the first time, he was offered a coupon. When entering the same site for the second time, no coupon was offered, but after removing the cookie and re-entering the site the coupon was offered again. Similarly, Dell customers have to trust that the company adjusts its prices in the customized sites as quickly and to the same extent as it does in the public Dell site.

The Supplier's Point of View

Benefits of Site Customization for Amazon and Dell

1. *Site-related benefits*. By showing products to customers that they may not even be aware of but that do match their interests, it is likely that Amazon increases the number of products customers browse and buy.

In the case of Dell, the opposite is probably true. By showing only approved items, Dell customers may spend less time in the site and consider fewer products. Given their different markets, both increased stickiness (Amazon) and decreased stickiness (Dell) are considered as positive.

2. *Transaction-related benefits*. By offering additional services through customization, which services even increase in value over time as the company learns more about its customers, Amazon customers are provided significant added value. These benefits may compensate not only for the fact that Amazon often does not offer the lowest price, but may even induce customers to buy more items. Dell Premier makes purchasing so much easier and faster for its customers' employees, that it is likely that Dell benefits with additional transactions. Also, given the initial investment that customers must make combined with the services Dell is offering (*e.g.,* asset management) customization probably has a positive effect, both on the share of budget Dell gains and on customer retention.

3. *Information-related benefits*. Customization provides both Amazon and Dell with much more information regarding the customer preferences, purchasing intervals and the products that customers consider and buy. Amazon can monitor the effectiveness of its recommendations in real time. Dell has information about approved configurations and available budgets. This information can be used to approach customers with better offerings and to improve sales forecasts, procurement planning and inventory management, respectively.

4. *Relationship-related benefits*. Both Amazon and Dell have designed their systems in such a way that customers derive most benefits in the long run when they make most purchases at Amazon or Dell. When Amazon is able to improve its recommendations over time, customers benefit more. The more details of their purchasing process (approved configurations, authorization levels, budgets, *etc.*) Dell customers

provide, the more they benefit from Dell Premier. Also, the benefits of PC asset management and customized reports are mostly reaped in the long run, thereby increasing the relationship-related benefits to Dell.

5. *Increased switching costs*. Amazon aims to increase switching costs by offering a purchase process that is more efficient (one-click buying) and effective (better recommendations leading to better purchases). However, most switching costs are only attitudinal and it remains relatively easy for customers to switch. In Dell's case, customer switching costs are increased considerably, due to the investments customers have to make to set up the service, time savings that are directly related to the number of purchases, the usefulness of functions depending upon non-switching (*e.g.,* the budget controls of Dell Premier cannot be used when customers spend part of their budget at other suppliers) and benefits that are accrued only over a longer period of time.

Costs of Site Customization for Amazon and Dell

1. *Initial investment*. Although both companies have not disclosed information about their monetary investments in customization, it is clear that the composition of the costs was different. At Amazon, most costs were initial and shrank; the marginal costs are very low. At Dell however, most costs (initially) were variable. By offering customization to the next customer, Dell had to cope with the additional costs of setting up the system. When customization turned out to be successful and increasing numbers of customers were added to the system, Dell invested heavily in automating the process to develop a customized site. That investment extended the Premier Service to more customers, countries and languages. As both companies are customization leaders in their fields, they had to develop the systems themselves.

2. *Knowledge*. Both companies relied on different knowledge to develop their customized sites. Amazon relied on sophisticated statistical

knowledge to produce useful recommendations. Dell relied on their understanding of the process leading to and following a computer purchase. Many of the services offered through Dell Premier are tasks that used to be performed by the customer and not by their hardware supplier. Detailed understanding of the process of organizational buying and asset management enabled Dell to provide services with substantial added value.

3. *Maintenance costs*. Both companies have tried to develop a system where maintenance is largely automated. When new data become available, the Amazon system can automatically adjust its recommendations. In other cases customers can maintain their data (*e.g.,* a new address or credit card). In case of new items the system needs only to know the product characteristics in order to determine which customers to offer this item to. At Dell, customers do most of the maintenance, *e.g.,* entering new approved configurations, budgets, employees, or authority levels.

4. *Failure to meet increased customer expectations*. As Amazon customers initially are not aware of and do not actively participate in site customization, their expectations are not high. When they engage in dialogues and provide Amazon with additional information, they expect Amazon to use this information effectively. However, in many cases the relationship between personal information (*e.g.,* the rating of a product) and customer benefits (improved recommendations) remains vague. At Dell, the relationship between entered customer data (*e.g.,* approved configurations) and expected customer benefits (only approved configurations are shown) is crystal clear.

CONCLUSIONS

Similar to mass customization, Website customization offers suppliers and customers the opportunity to develop tailor-made digital services and communication. Site customization is a rich concept but comes with benefits and costs to both customers and suppliers. In this chapter, we identified the value drivers for customers and suppliers and used that model to analyze the customization efforts of Amazon and Dell.

The analyses show many similarities in the customization strategies of both companies, but also distinct differences (see Table **3**). Both companies clearly manage to deliver substantial added value to customers using the customized features. Customers save time, get better information and see more relevant alternative products. Both companies aim to minimize the marginal costs of setting up site customization by automating the process of setting up the service to a very high extent (Dell) or fully (Amazon). They both offer a range of customized features that can be implemented independently. This implies that customers can start using the features at a low level, requiring minimal user input. After successfully testing the waters, customers can expand the use of customization. Both companies realize a broad range of advantages through customization, including improved transactions, information and customer relationships. Finally, in both cases trust plays a major role. When customers do not intend to have a long lasting relationship with their supplier site customization is unlikely to be successful (for both parties).

Table 3: Overview of How Amazon and Dell Deal With the Various Value Drivers of Site Customization and the Extent to Which Both Companies Have Implemented a Similar Strategy.

	Amazon	**Dell**	**Similar?**
Customer benefits			
Information and time savings	More relevant and faster process	More relevant and faster process	H
Useful recommendations	Items according interests	Items according pre-approval	L
Increased sense of control	Free to use/turn off features	Free to use features	H
Sense of esteem	Highly personalized site	Highly personalized site	H
Customer costs			
Time	Initially none	Initially considerable	L
Cognitive efforts	Only password	High level of formalization	L
Privacy violation	Collection and use of personal data	Security due to confidentiality	L
Exploitation by supplier	High trust needed	High trust needed	H

Table 3: cont….

Supplier benefits			
Site-related benefits	Increased stickiness	Decreased stickiness	L
Transaction-related benefits	Increased sales	Increased sales	H
Information-related benefits	Customer preferences and purchases	Customer preferences and purchases	H
Relationship-related benefits	Most customer benefits in long run	Most customer benefits in long run	H
Increased switching costs	To some extent	To a large extent	M
Supplier costs			
Initial investment	Substantial, mainly sunk	Substantial, both sunk and marginal	M
Knowledge	Statistics	Organizational purchase process	L
Maintenance costs	Much automated/self-service	Much automated/self-service	H
Failure to meet increased customer expectations	Expectations managed partly	Expectations managed clearly	M

Note. H: highly; M: medium; L: low level of similarity.

There are also marked differences between both companies. As Amazon is active in a low-margin business, it automated the entire customization process, thereby bringing down the marginal costs to (almost) zero. At Dell, employees are still involved in setting up the system. For this reason, Amazon is able to offer its service to any customer, whereas Dell only offers customization to its larger customers. Furthermore, the initial investment customers have to make is quite high at Dell, whereas it is virtually nonexistent at Amazon. Another important difference is that Amazon customers gain by making better purchase decisions, whereas Dell customers derive most value from pre- and post-purchase activities. Amazon automates a good salesperson; Dell automates processes that used to be performed by their customers and not by IT suppliers. Finally, both systems are based on different knowledge; contrast the advanced statistical knowledge of Amazon to the detailed organizational procurement process knowledge of Dell.

This study identified the value drivers for both customers and suppliers of site customization. By doing so, we developed a tool that both managers and researchers can use to investigate (possible) applications of site customization (see Table **3**).

Managers can use the checklists to determine how likely it is that site customization will lead to a pay off that justifies the required investments. Managers can also apply the customer cost and benefit analysis to determine how likely it is that customers will actually participate in site customization. The list of customer benefits can be useful when managers need to persuade customers to use the offered site customization features. The list of customer costs can be useful to minimize the threshold that customers face when considering the use of site customization. Finally, the analysis of both case studies can serve as a useful starting point for managers to develop ideas about potential customization applications.

Researchers can use the components of costs and benefits to estimate the effectiveness of site customization in various situations. This kind of research will provide deeper insight into the situational requirements of successful applications of site customization. Such studies can focus on suppliers (for which suppliers is customization a fruitful option?), customers (which customers are likely to adopt site customization?), relationships (what are the relational requirements for successful application of site customization?) and applications of site customization (what kind of applications are most successful?). These studies can be cross-sectional, to study site customization in various situations, as well as longitudinal, to study the evolving use of site customization over time.

Site customization appears to be the classical case of "win–win". Only in cases when both a supplier and a customer perceive positive net benefits are they willing to engage in site customization. From the work of Nobel prize winner Kahneman and his colleague, the late Amost Tversky (1979), we know that customers weigh losses (costs) much heavier that gains (benefits), implying that companies have to carefully manage the costs and benefits of site customization in order to develop successful applications of a potentially rich concept.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Anderson, C. (2006). The long tail: Why the future of business is selling less of more. New York, NY: Hyperion.

Angehrn, A. A. & Meyer, J. F. (1997). Developing mature Internet strategies. Information Systems Management, 14(3), 37–43.

Ansari, A. & Mela, C. F. (2003). E-customization. Journal of Marketing Research, 40(2), 131–145.

Bardakci, A. & Whitelock, J. (2003). Mass-customization in marketing: The consumer perspective. Journal of Consumer Marketing, 20(5), 463–479.

Bizrate. (2000). 75 Percent of online buyers abandon shopping carts. Press release Bizrate.com, October 23.

Bolton, R. N. (1998). A dynamic model of the duration of the customer's relationship with continuous service provider: The role of satisfaction. Marketing Science, 17(1), 5–65.

Bunduchi, R. (2005). Business relationships in Internet-based electronic markets: The role of goodwill trust and transaction costs. Information Systems Journal, 15(4), 321–341.

Burke, R. R. (2002). Technology and the customer interface: What customers want in the physical and virtual store. Journal of the Academy of Marketing Science, 30, 411–432.

Da Silveira, G., Borenstein, D., & Fogliatto, F. S. (2001). Mass customization: Literature review and research directions. International Journal of Production Economics, 72, 1–13.

Dann, S. & Dann, S. (2004). Strategic Internet marketing 2.0. Milton QLD: Wiley and Sons Australia.

eMarketer. (2004). Abandon shopping cart! Press release eMarketer.com, January 14.

Festa, P. (2003). Report slams Web personalization. CNET News.com, October 14.

Franke, N. & Piller, F. (2004). Value creation by toolkits for user innovation and design: The case of the watch market. Journal of Product Innovation Management, 21, 401–415.

Franke, N., Keinz, P., & Steger, C. J. (2009). Testing the value of customization: When do customers really prefer products tailored to their preferences? Journal of Marketing, 73(5), 103–121.

Godek, J., Diehl, K., Zauberman, G., Hamilton, R. W., Sivakumaran, B., Yates, J. F., & Yoon, Y. (2002). Special session summary: Personalization and customization: Implications for customer decision making and behavior. Advances in Customer Research, 29, 155–157.

Grover, V. & Saeed, K. A. (2004). Strategic orientation and performance of Internet-based businesses. Information Systems Journal, 14(1), 23–42.

Hanson, W. (2000). Principles of Internet marketing. Cincinnati, OH: South Western College Publishing.

Hart, C. W. (1995). Mass customization: Conceptual underpinnings, opportunities and limits. International Journal of Service Industry Management, 6, 36–45.

Hoffman, D. L., Novak, T. P., & Chatterjee, P. (1995). Commercial scenarios for the web: Opportunities and challenges. Journal of Computer-Mediated Communications, 1, 1–20.

Huang, E. Y. & Lin, C. Y. (2005). Customer-oriented financial service personalization. Industrial Management & Data Systems, 105(1), 26–44.

Huizingh, K. R. E. (2002). Towards succesful E-business strategies: A hierarchy of three management models. Journal of Marketing Management, 18(7/8), 721–748.

Jayawardhena, C. & Wright, L. T. (2009). An empirical investigation into e-shopping excitement: antecedents and effects. European Journal of Marketing, 43(9/10), 1171–1187.

Johnson, E. J., Bellman, S., & Lohse, G. L. (2002). Cognitive lock-in and the power law of practice. Working Chapter Colombia University.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, XLVII, 263–291.

Kamali, N. & Loker, S. (2003). Mass customization: Online customer involvement in product design. Journal of Computer-Mediated Communications, 7, 1–21.

Kasanoff, B. (2002). Making it personal: How to profit from personalization without invading privacy. New York, NY: Perseus Books Group.

Kiang, M. Y., Raghu, T. S., & Shang, K. H. (2000). Marketing on the Internet – Who can benefit from online marketing approach? Decision Support Systems, 27, 383–393.

Kotler, P., Armstrong, G., Saunders, J., & Wong, V. (2003). Principles of marketing (3rd European Edition ed.). Harlow, UK: Pearson Education Limited.

Liao, C., To, P.-L., & Shih, M.-L. (2006). Website practices: A comparison between the top 1000 companies in the US and Taiwan. International Journal of Information Management, 26(3), 196–211.

Mick, D. G. & Fournier, S. (1998). Paradoxes of technology: Consumer cognizance, emotions and coping strategies. The Journal of Consumer Research, 25(2), 123–143.

Murthi, B. P. S. & Sarkar, S. (2003). The role of management sciences in research on personalization. Management Science, 49(10), 1344–1363.

Nilson, T. H. (2002). Customize the brand: Make it more desirable and profitable. Hoboken, NJ: Wiley and Sons, Inc.

Nunes, P. F. & Kambil, A. (, 2001). Personalization? No thanks. Harvard Business Review, 79, 32–34.

Peppers, D. & Rogers, M. (1999). Is your company ready for one-to-one marketing? Harvard Business Review, 77(1), 151–161.

Porter, M. E. (March, 2001). Strategy and the Internet. Harvard Business Review, 63–78.

Reibstein, D. J. (2002). What attracts customers to online stores and what keeps them coming back? Journal of the Academy of Marketing Science, 30, 465–473.

Reichheld, F. F. & Schefter, P. (July–August, 2000). E-loyalty your secret weapon on the web. Harvard Business Review, 105–113.

Rowley, J. E. (2002). Reflections on customer knowledge management in e-business. Qualitative Market Research, 5, 268–280.

Schafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. Data Mining and Knowledge Discovery, 5, 115–153.

Schoenbachler, D. D. & Gordon, G. L. (2002). Trust and customer willingness to provide information in database-driven relationship marketing. Journal of Interactive Marketing, 16(3), 2–16.

Schultze, U. & Orlikowski, W. J. (2004). A practice perspective on technology-mediated network relations: The use of Internet-based self-serve technologies. Information Systems Research, 15(1), 87–106.

Srinivasan, S. S. anderson, R., & Ponnavolu, K. (2002). Customer loyalty in e-commerce: An exploration of its antecedents and consequences. Journal of Retailing, 78, 41–50.

Surmacz, J. (2003). The personalization myth. CIO.

Swaminathan, J. M. (2001). Enabling customization using standardized operations. California Management Review, 43, 125–135.

Taylor, R., Terwiesch, C., & Ulrich, K. T. (2005). Principles for user design of customized products. California Management Review, 47(4), 68–85.

Teerling, M. L. & Huizingh, E. K. R. E. (2006). Exploring the concept of website customization: Applications and antecedents. University of Groningen, SOM Research Report, 06F07, 32p.

Teo, T. S. H. (2005). Usage and effectiveness of online marketing tools among Business-to-Consumer (B2C) firms in Singapore. International Journal of Information Management, 25(3), 203–213.

Thompson, M. (1999). Registered visitors are a portal's best friend. The Industry Standard. June 7, 1999.

Vrechopoulos, A., Katherine, P., Pramataris, C., Doukidis, G., & Lekakos, G. (2003). An Internet retailing data framework for supporting consumers and business processes. Information Systems Journal, 13(4), 353–373.

Williamson, O. E. (1975). Markets and hierarchies. Englewood Cliffs, NJ: Prentice Hall.

Wind, J. & Rangaswamy, A. (2001). Customerization: The next revolution in mass customization. Journal of Interactive Marketing, 15, 13–32.

Wolfinbarger, M. & Gilly, M. C. (2003). ETailQ: Dimensionalizing, measuring and predicting etail quality. Journal of Retailing, 79, 183–198.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality and value: A means-end model and synthesis of evidence. Journal of Marketing, 52(3), 2–22.

PART VI

MIND GENOMICS®: HOLISTIC UNDERSTANDING OF CONSUMERS

CHAPTER 19

Introduction to Mind Genomics®

Alex Gofman*

Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: It's almost impossible to imagine the scale, cost and labor Involved in the human genome mapping projects. A similar scale effort to map the human *mind* would be even more complex, as people's mind-sets are more individual. In this chapter, we discuss the new science called Mind Genomics® and how it can be employed in a number of subject areas.

Keywords: Human genome, cognition, mind-sets, consumer preferences, perceptions.

INTRODUCTION

It's difficult, if not virtually impossible, not to appreciate the almost billion-dollar, 13-year multinational effort to map the human genome, a project whose resultant amount of information could be compared with a book of over one billion words long (a dollar a page—sounds like a deal), bound in 5000 volumes, each 300 pages long. However, a suggestion to systematically map the human *mind,* even if limited to consumer perceptions and preferences, could easily dwarf it. With 100 billion neurons, the task of understanding the brain is immensely more difficult. Despite the years of hard work of thousands of cognitive and neuroscientists around the world, we're just scratching the surface (Greenfield, 2000). Unlike DNA code, which is virtually the same for every person in the world and omnipresent in practically every cell of our bodies, people's mind-sets are more individual and the challenge is to see what ideas are shared by what proportion of the population. So, multiply the amount of information of people's preferences by the population size. Sound daunting?

Mind Genomics® is a new science introduced in 2005 by a group of academics and practitioners. It aims to systematically map consumers' perceptions and preferences. Unlike the genome project, which became usable mostly after the completion, understanding the mental genome of the population by collecting and

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

databasing structured information, topic by topic, market by market, Mind Genomics® is usable as soon as a fair representation of the subject is established.

The business objective is to get deep understanding of consumers and to intercept trends, take advantage of this knowledge with novel offerings, and, hopefully, succeed by being there first. The realities of business nowadays dictate that knowledge needs to be available in "Google time," at the press of a button, pre-digested and ready to be used in nice, neat buckets, all done with little budgets. Today's marketers and developers need off-the-shelf, almost shrink-wrapped, systematized knowledge—organized insights about the customer's mind in specific topic areas, to guide development on one hand and messaging on the other.

Mapping the mental genome is modeled on the emerging science of genomics and the technology of informatics. The goal is to better understand how people react to ideas in a formal and structured way, using the principles of stimulus–response (from experimental psychology), conjoint analysis (from consumer research and statistics), Internet-based testing (from marketing research), and multiple tests to identify patterns of mind-sets (patterned after genomics). This formal approach can then be used to construct new, innovative ideas in business.

Instead of having a database of activities summarizing what information is known from who buys what, economic trends, and so on, Mind Genomics® creates a ready-to-use database about the set of ideas in the minds of consumers, to be used for advertising, product development, and merchandising. Or, it is simply used to understand the way people think about specific topics.

What is the nature of such a database? One can begin creating many of these databases fairly simply to ultimately develop a marketer's library. Such a library could be updated on a regular basis, comprising the collections of the mind-set related to different topics. Each database might pertain to one topic, such as shopping, insurance, the fast food experience, or the automobile experience. Each study in the database would deal with one specific aspect. For example, in the case of automobiles, we would have separate studies concerning comparison shopping on the Web, layouts of the automobile showroom, test drives, financial payments, car design, car advertising on television, and so on. Each of these separate studies would, in turn, comprise experimentally designed vignettes to understand the algebra of the customer's mind, as well as extensive self-profiling classifications. With these databases available, there might be a simple pay-as-you-go digital, searchable library that can quickly reveal what preferences and ideas consumers currently hold in their minds on different topics. Online tools could be created to invent new ideas by recombining old ideas and novel, inputting ideas into new mixtures using the recombinant genomics approach (combinatorial innovation).

Aside from the work of the academics and practitioners (which can be explored at http://www.mindgenomics.org), there are a number of other Internet-based approaches which are starting to track (although not necessarily systematically map) the mind of the consumer, such as Buzz Metrics and Google Trends. Google Trends (http://www.google.com/trends) dynamically tracks the most searched information on the Web and is available free of charge. By picking up such fuzzy signals of possible emerging consumer interests and analyzing the mental genome database of the related existing consumer perceptions and preferences, a researcher could spot a promising next "big thing" and give it an immediate spin to generate new product ideas.

Researchers at Indiana and Northeastern Universities, along with companies such as McCormick, Symrise, Guardian Life Insurance, and Wild Flavors, are already implementing the approach using Mind Genomics® data as the initial seeds for ideation, brainstorming and multicultural marketing, and building other applications not yet imagined.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This introduction is based on Gofman, A. (2009). Mind mapping. *Research World, 8*, 26–28.

REFERENCE

Greenfield, S. (2000). The private life of the brain: Emotions, consciousness, and the secret of the self. Hoboken, NJ: Wiley and Sons, Inc.

CHAPTER 20

Mind Genomics®: A Systematic Consumer Research

Howard R. Moskowitz[1,*], Alex Gofman[1], Jacqueline Beckley[2] and Hollis Ashman[3]

[1]Moskowitz Jacobs Inc., White Plains, New York, USA; [2]The Understanding and Insight Group, Inc., Denville, New Jersey, USA and [3]Muse LLC Powell, Ohio, USA

Abstract: This chapter presents our vision for a new science, modeled on the emerging science of genomics and the technology of informatics. Our goal in this new science is to better understand how people react to ideas in a formal and structured way, using the principles of stimulus–response (from experimental psychology), conjoint analysis (from consumer research and statistics), Internet-based testing (from marketing research) and multiple tests to identify patterns of mind-sets (patterned after genomics). We show how this formal approach constructs new, innovative ideas in business. We demonstrate the approach using the development of new ideas for an electronic color palette for cosmetic products that are to be used by consumers.

Keywords: Conjoint analysis, consumer research, structured experimentation with consumers.

INTRODUCTION

During the past several decades, the emergence of computation as a major driver of scientific prowess has accelerated. When first developed in the 1940s, much of the statistical computation was done either manually by so-called computers (*i.e.* individuals who did the computation) or by sorting machines such as the Hollerith card sorting machine. At that time, the use of statistics was relatively minor, confined to those types of statistical tests that could be executed easily in the field or in the laboratory. The notion of larger-scaled analyses using statistical methods was acceptable, but more often in the realm of fantasy than fact. Indeed, the senior author has fond (and occasionally not-so-fond) memories of manually analyzing data from studies with a professor at Queens College, New York. The data, collected by Professor Louis M. Herman in the late 1950s at Wright Patterson Air

***Address correspondence to Howard R. Maskowitz:** Maskowitz Jacobs Inc., White Plains, New York, USA; Tel: 914-421-7408; Fax: 914-428-8364; Email: mjihrm@sprynet.com

Alex Gofman and Howard R. Moskowitz (Eds)

Force Base, were analyzed during a four-month internship in 1962 by the senior author and Jerry Weiss, both undergraduate students in psychology. The manual analysis of what today would be considered a simple three-way analysis of variance, took approximately three months, using the MonroeMatic® and Friden® calculators and an assortment of scratch paper to record intermediate results.

During those years, the notion of genomics was becoming of increasing interest, but the thought that someday such thinking could generate easy-to-execute studies for the Genome project was far away (Collins, Green, Guttmacher and Guyer, 2003). The notion was even further in the future that recombining ideas rather than genes using statistical design could be a reality and indeed as will be shown, a very simple reality. It would be the confluence of statistics and genomics, especially the simplicity of executing work in both, that would become the inspiration for the science described here (Systat, 1997; Van Ommen and Stierum, 2002; Watkins and German, 2002).

It was quite another thing to use statistical design to understand the consumer mind and to move such understanding to creating ideas for product business in the commercial world. However, such applications of experimental design were soon in coming, driven by a scientific renaissance after Sputnik. The psychological sciences benefited as much as the physical and biological sciences. In the early 1960s, a frenzy of newly funded research efforts concentrated on better understanding the underpinnings of psychological measurement (Suppes and Zinnes, 1963). One of the most important developments of this period was conjoint measurement, at first a product of mathematical psychologists seeking to better understand the underpinnings of measurement (Luce and Tukey, 1964), but destined to grow into a major intellectual development stream that would spark radical developments in consumer research and business (Green and Rao, 1971; Green and Srinivasan, 1990; Moskowitz, Porretta and Silcher, 2005c; Wittink, Vriens and Burhenne, 1994).

Conjoint measurement, the basic quantitative structure underlying the science proposed in this chapter, can be reduced to a simple descriptive statement, namely the use of experimental design to understand reactions to ideas by measuring reactions to mixtures of ideas. Conjoint measurement uses experimental design, mixing together small components (or *idealets*, if such a word were to be used), generating combinations, acquiring subjective responses to those combinations and then deducing what components drive the reactions (Box, Hunter and Hunter, 1978).

THE FOUNDATION OF THE APPROACH

The science of Mind Genomics® began in the last part of the 1990s, as the notion of archiving utility values went from dream to reality. Prior to the development of conjoint analysis using dummy variable modeling, which allowed for meaningful values of element utilities, most conjoint analysis used so-called complete concepts. Each concept comprised exactly one element from each of the available silos. Thus, because of statistical multicollinearity, analysis of the data could not generate absolute values for the utilities. That is, the desire of users of conjoint measurement to work with complete concepts meant that statistical regression analyses of the data developed relative utilities.

Differences between the utility values were meaningful within a silo, but not the utility values themselves. Nor, in fact, could utility values be compared across silos; indeed, they could not be compared across studies. The use of incomplete concepts and modeling by dummy variable regression generated absolute values of utilities, first making the research easier, but more importantly, creating the foundations of a valid science. As long as conjoint analysis used (and continues to use) complete concepts with the inevitable multicollinearity that ensues, it will be impossible for the utilities to have real meaning. There may be conference after conference, article after article dealing with these issues of making otherwise non-meaningful utilities meaningful, but they will only be statistical exercises, not the foundations of a science.

CREATING THE NEW SCIENCE—SETS OF LINKED STUDIES GENERATING A SEARCHABLE DATABASE

Conjoint analysis, the output of efforts in mathematical psychology and a key tool to understanding and prioritizing features in marketing, now became the basis of this new science. The new science itself—Mind Genomics®—creates a corpus of knowledge about how people respond to the components of a complex stimulus. The objective of this science is to create databases about what features in product descriptions or situational "vignettes" are important. At the surface level, the science quantifies what is important. At a deeper level, the science creates a body of knowledge that reveals how people think about different topics, working from responses to complex vignettes downward to more fine grained granularity as to how specific components contribute.

First attempts at creating this corpus of knowledge can be seen by two initiatives to create databases of concept ideas using conjoint analysis and attempting to formalize the knowledge structure. These are the It!® databases and the InnovAidOnline™ initiative.

THE IT!® DATABASES

When the first steps toward this science were taken in early 2000, the objective was simply to understand what features of products or what specific communications and brand names make products craveable. The strategy was to work with a number of different products, not just one product alone. What transformed the approach from a research project to a science was the creation of linked databases, the structured approach and the enormous potential to understand new, hitherto unexpected aspects of consumer responses, either from patterns of the individual utilities alone within a product study, across different products, or integrating the conjoint portion of a study with the self-profiling or classification portion of the same study.

The early efforts created the science by developing a so-called mega database of 30 related studies. Those early efforts were followed by updated mega databases for craving (called the Crave It!® series), as well as other large-scale databases for beverages (Drink It!®) and a number of other different databases. As can be seen below, the databases span a range from foods to lifestyles.

It!® Databases that have been created and venues where the results have been presented or published:

1. **Foods (Crave It!®)**

   a. Crave It!® USA (Moskowitz, Ashman, Gillette and Adams, 2002a)

b. Eurocrave (Aarts, Paulus, Beckley and Moskowitz, 2002)

c. Eurocrave (Luckow, Aarts and Moskowitz, 2003)

d. Teen Crave It!® (Ashman, Beckley, Adams and Mascuch, 2002)

e. Beckley, Gillette and Marketo (2002)

f. Beckley, Ashman, Maier and Moskowitz (2004a, 2004b)

g. Moskowitz, Beckley and Adams (2002b)

h. Moskowitz, Silcher, Beckley, Minkus-McKenna and Mascuch (2005d)

i. Beckley, Moskowitz and Minkus-McKenna (2004c)

j. Krieger, Ashman and Mascuch (2002)

2. Beverages (Drink It!®)

a. Hughson, Ashman, de la Huerga and Moskowitz (2004)

b. Poskanzer and Ashman (2003)

3. Fast food experience (It's! Convenient)

a. Ashman and Beckley (2004)

4. Healthful products (Healthy You!)

a. Zawel (2002)

b. Luckow, Moskowitz, Beckley, Hirsch and Genchi (2005)

5. Anxiety (Deal With It!®)

a. Ashman, Teich and Moskowitz (2004)

b. Beckley and Ashman (2004)

c. Moskowitz, Itty, Ewald and Beckley (2004)

6. The customer shopping experience (Buy It!®)

a. Himmelstein, Ashman, Moskowitz, Minkus-McKenna and Rabino (2004)

b. Moskowitz and Ashman (2003)

c. Minkus-McKenna, Ashman and Moskowitz (2004)

d. Ashman, Rabino, Minkus-McKenna and Moskowitz (2003)

In addition to the published databases, there are other databases in the effort, including those dealing with insurance (Protect It!®) and not-for-profit topics (Give It!®).

The reason that the It!® databases may be construed to be the first contributions to this new science is the combination of systematics (attempts to catalog mind-sets using conjoint analysis), rules (attempts to understand general trends), applications (attempts to use the results to create new ideas) and testable predictions (attempts to predict success of products or new trends from the conjoint results, which quantify level of consumer interest in the idea or set of ideas).

EXEMPLIFYING ASPECTS OF THE NEW GENOMICS OR COMBINATORIAL SCIENCE USING COSMETICS

Case histories are often a good way to illustrate new ideas because the case history takes the approach from early-stage thinking to a working example with testable results. We will illustrate the Mind Genomics® approach using data from a small demonstration with cosmetics. Cosmetics constitute a mainstay of consumer packaged goods. With the fierce competitive environment worldwide, with the unpredictability of styles and fads and with the expense of marketing and merchandising products in this ever-changing environment, anything that helps better create new ideas will be welcome to the enterprising business person.

The objective of the case history was to synthesize a new product—the electronic color palette—that could help a user identify the optimal colors for different cosmetic products. The ingoing hypothesis was that a synthesis of cosmetics and electronics provides a new thrust for companies looking to differentiate in a competitive market.

The first three studies dealt with lip cream, eye shadow and a skin color sensor, respectively. The main objective of the first three studies was to develop a small database of features that interest consumers. The fourth study, dealt with below, synthesized a new-to-the-world combination of features, by splicing together ideas from different products into a new "whole". The two-phase exercise demonstrates both the approach to developing the database and the synthesis using genomics-inspired recombination of elements (Moskowitz, 2001).

Step 1: The Silos and Elements

The elements are at the heart of the study. For the three studies, each element could be classified as belonging to one of six silos, as shown in Table **1**. The same set of six silos applied to each of the three studies.

Table 1: The Six Silos for Each Product.

Silo	Topic of Silo
A	What is it? (Product definition)
B	Who will use it, for whom is it designed?
C	Mode and ease of use
E	Features and benefits
E	Additional features (add-ons, special characteristics)
F	Where is it sold, how it is merchandised, where do you find it?

Step 2: Experimental Design

Each of the studies generated a unique set of 400 experimental designs, all of which were permutations of the same basic design (Gofman and Moskowitz, 2010). That basic design comprised 48 combinations of the 36 elements, such that each element appeared independently of every other element, an equal number of times, against randomized combinations. This strategy ensures that no particular combination of two elements can unduly influence the results. Thus, if a pair of elements synergize so that the combination does much better or worse than expected, such an interacting pair appears only some of the time across the many combinations and thus its impact on the data is minimal. With the approximate 100–130 respondents participating in a study, the permutation strategy ensured that each person would be presented with a unique set of combinations, although the person would always test all of the individual elements.

Step 3: Field Execution

A total of 6,000 invitations were sent to individuals who had previously indicated that they would like to participate in these types of studies. The invitation presented all three studies, from which the participant could choose one that was most interesting. The individuals were given a relatively narrow time slot in which to participate (4 days), which decreased the standard rate of 9% in these types of studies to a slightly lower response rate of 6%. The color sensor study generated 117 completes of 175 log-ins, the lip gloss study generated 127 completes of 171 log-ins and the eye shadow study generated 125 of 193 log-ins, respectively.

Step 4: Basic Results

Each participant evaluated 48 different combinations for eye shadow, lip cream, or color sensor, respectively, using an anchored 1–9 rating scale. The ratings for each participant are converted into a binary response, with original ratings of 1–6 converted to the value “0,” and ratings 7–9 converted to the value “100”. The conversion changes the focus from the intensity of the participant’s feeling about the product to membership in the class of “concept acceptors” (7–9 or 100), or membership in the class of “concept rejecters” (1–6 or 0). Such focus on membership diminishes some of the metric information in the data, but conforms to the conventions of consumer research, which is interested in group membership. It’s important to keep in mind that with 48 concepts presented to each person, every individual may for one concept be considered an "acceptor" because of the rating of 7–9 and yet for another concept be considered a "rejecter" because of the rating of 1–6. That is, acceptance/rejection is contingent on response to an individual concept, not to the entire product.

The data for each participant are subject to regression modeling, which is perfectly valid for these types of results since each individual participant evaluated 48 concepts specifically designed to be analyzed by regression modeling. The experimental design ensures that all 36 elements are statistically independent of each other.

The results of the study, from the entire set of participants, appear in Table **2** (see previous chapters for explanations of the modeling). The results can be interpreted quite simply as follows:

Table 2: Performance (Utility Value) of the 36 Elements and the Additive Constant for the Three Products (Partial List).

Eye Shadow		**Lip Cream**		**Color Sensor**	
Base size (number of participants in the study who completed the interview)					
	125		127		117
Additive Constant (basic interest in idea without elements)					
	43		61		48
Silo A—What is it? (Product definition)					
A dazzling collection of six perfectly harmonized eye shadows.bring out the best in your eyes	8	Long-lasting color and shine in a compact portable palette	4	A color detector designed to mix and match colors	0
A racy palette with six dramatic shades in one slim compact	6	Moisturizing, long-lasting lip color…perfect for any skin tone	2	Find the perfect match with a personal color detector.superior color capabilities right at your fingertips	–1
Six perfectly coordinated shades at your disposal.to create an endless number of stunning looks	5	A sassy array of lip cream color…the quickest way to brighten up your face	1	A high-resolution color detector with shade matching capabilities	–1
Six-color combo.designed to enhance and bring out your natural eye color	3	Ravishing lip cream that is easy to layer and blend…make a colorful impression	–4	A pocket-sized device detects colors that are best matched together	–3

Table 2: cont….

Silo B—Who will use it, for whom is it designed?					
Perfect to wear day and night…perfect for any occasion	2	For the woman who's always on the run…easy to use wherever you are	0	Ultra-reliable.state-of-the-art technology for scientists, production managers and other professionals	1
For the girl who can't commit to one eye shadow	–5	Potent color and luxurious feel…for the daring and seductive woman	–2	A pleasing gift for the artist and technology fanatic alike	–2
Perfect for the refined woman wanting to make an elegant statement at the social event of the year	–6	Ideal for the refined woman who wants to make an elegant statement at the social event of the year	–8	For the technology-savvy individual looking for a new toy	–4
Silo C—Mode and ease of use					
It's easy to blend colors to achieve the desired look	6	Won't smudge, run, or transfer…so you can eat, drink and kiss without having to reapply	7	Easy to use.great for any color needs	1
Easy to apply, easy to remove…beauty in a matter of seconds	6	Mix, layer, lighten, or intensify.achieve the perfect shade for every occasion	4	Small and lightweight.fits perfectly into your back pocket	0
Silo D—Features and benefits					
The 18-hour long-wear eye shadow that won't smudge, run, or fade	11	Satin feel and finish…for creamy, moist-looking lips	6	Small light beams can sense the difference between matte and glossy and detect the finest nuances in color	5
Shimmery color kisses lids…adds a bit of glamour to your look	9	Hydrating gel drenches lips.lips feel moisturized even after you take it off	4	Accurate color differentiation.match colors as precisely as possible	2
Made with hypoallergenic ingredients for contact wearers and sensitive eyes…dermatologist and ophthalmologist approved	8	Sheer color with the perfect hint of shimmer and shine	3	Connects to your computer, PDA and a variety of other devices.the perfect color companion	1
Rich, long-wearing and crease-free…your eyes deserve the best	8	Silky soft lip cream…set the stage for intrigue	0	High resolution and accuracy.easily and precisely distinguish between shades of colors	0
Silo E—Additional features (add-ons, special characteristics)					
Enhanced with alpha hydroxy and fruit acids…improve skin's texture	4	Formulated with retinol to visibly reduce lip lines.and collagen for visibly fuller lips	1	Additional removable memory chip stores shades and hues.so you can keep all the colors you create	2

Table 2: cont….

With a unique blend of oil-free moisturizers…so your eyelids feel silky smooth	4	Enhanced with SPF 15…gives your lips the utmost protection	–1	So reliable, it comes with an extended 10-year warranty	0
With a bonus brush for error-proof application	3	Enriched with vitamin E and aloe.increase wearability and keep the color true	–2	With an additional car charger.so you can charge and go, always there when you need it	–1
Silo F—Where is it sold, how it is merchandised, where do you find it?					
Available at your neighborhood drug store	11	Available in your local drug store	–3	Find it in the electronics section of department stores nationwide	0
Buy it directly from the manufacturer's online Website	–13	Available in beauty retail stores like Sephora	–23	Available in your neighborhood technology and electronics dealer	–5
Available in beauty retail stores like Sephora	–14	Buy it directly from the manufacturer's online Website	–25	Buy it directly from the manufacturer's online Website	–10
Purchase through a mail-order catalog and have it delivered right to your door	–14	Purchase through a mail-order catalog and have it delivered right to your door	–25	Purchase through a mail-order catalog and have it delivered right to your door	–14
Purchase with the aid of a personal sales representative in the comfort of your home	–27	Purchase with the aid of a personal sales representative in the comfort of your home	–35	Purchase with the aid of a personal sales representative in the comfort of your home	–17

Note. All data come from the total panel of participants for each study. Elements are sorted within a silo from best performing to worst performing.

The Additive Constant in Light of the Nature of the Study Participants. Keep in mind that these participants are self-selecting, because they know from the invitation that the study would deal with women's health and beauty aid products. We can compare the additive constant of eye shadow to the additive constant for lip cream (61) and to the color sensor (48). Lip cream is more interesting. Part of the ingoing "vision" of Mind Genomics® is simply to obtain normative databases for such product areas.

Explicating the Element Utilities. The 36 individual elements, falling into the six silos, give us another sense of the product ideas. Let's first look at eye shadow. There are some very strong performing elements, but not many. Recall the definition of the element as the conditional probability of a participant being interested in the product (*i.e.* switching from a rating of 1–6 denoting not interested, to a rating of 7–9 denoting interested). A strong element is: *The 18-hour long-wear eye shadow that won't smudge, run, or fade*. Another strong element is: *Available at your neighborhood drug store*. Both elements have utility values of +11, which, from previous studies, would suggest a very strong performing idea. Indeed, with so many elements mixed and matched against different backgrounds, it is virtually impossible for a weak performing element to do well by "accident". There are too many variations.

Not Every Idea Performs Well. Silo B, which deals with "ease of use" and "who will use it" clearly shows some poor performing ideas with negative utilities, such as *Perfect for the refined woman who wants to make an elegant statement at the social event of the year*. This element has a utility of –6, meaning that when it is added to the concept the interest goes down.

Using Normative Data or Benchmark Results. Normative data from these types of studies suggest that the really strong elements perform 15 or higher, strong elements perform 10 or higher and good but not great elements perform 6 or higher. For the most part, the elements only perform modestly (around 0–5). Such modest performance for the total panel is to be expected if the elements attract some groups of individuals but repel other groups. We will see this type of segmentation into some groups that like and other groups that dislike the elements in the next section.

Step 5: Looking for Key Segments or "Mind-Sets" in a World Awash With Choice

The approach of Mind Genomics® to segmentation comes from a worldview different from the more conventional marketing approaches. The ingoing assumption of Mind Genomics® is that there exist segments in the population, much as the traditional marketer might believe. However, these segments may not be general. They do not cross over different categories. These segments manifest themselves simply as patterns of responses to concepts. Those three assertions about segments radically differ from the overarching approaches implicitly (and often explicitly) promoted by marketers.

Segments emerge from standard statistical analysis of the patterns of utilities at the level of the individual participant. The utilities used for segmentation are the so-called *persuasion utilities*, which are the regression coefficients for the different elements (but not the additive constant) estimated at an individual by individual basis *before* any binary transformation. The segmentation is accomplished by simple, well-accepted methods, such as first defining the distance between pairs of participants by a distance measure (*e.g.,* the value (1—Pearson correlation)) and then using that distance measure to put people into different groups such that people in the same group or segment are "close to each other," and people in different segments are "far away from each other".

The reality of these segments is in the fact that they make sense, that they emerge in similar ways time after time in different studies almost like archetypes and that they can be used to create product ideas along with more powerful communications. The segments have to be understood by the pattern of the ideas that they comprise. We have to stand back to see the nature of the segment itself. Most of the time we will see the segments emerge simply as a set of related elements, scored well by a subset of individuals in a study.

The following section investigates three different health and beauty aid products (lip cream, eye shadow and skin color sensor). We segmented the participants in each study into three groups. The segmentation or clustering is a formal statistical operation. Now with these data let us see which specific elements perform well.

The data for each of the studies were separately analyzed, with participants put into segments based on the pattern of their individual utilities. We looked at the three-segment solution for each product, to see whether we could create three general segments, transcending a specific product type. This attempt at creating super segments somewhat stretches the meaning a bit for each segment, but the approach allows us to treat the data in a more direct fashion. The super segmentation is not necessary, simply convenient. The three emerging general segments appearing in Table **3** are:

1. Segment 1—interested in short messaging, basic benefits, best performing elements only have modest utility;

2. Segment 2—interested in extra features, "techie";

3. Segment 3—what can be accomplished with the technology.

Table 3: Winning Elements for Three Super-Segments Developed from the Cosmetic Data.

Study	Super segment and element	Utility
Segment 1—Interested in short messaging, basic benefits, best-performing elements only have modest utility		
Eye	Available at your neighborhood drug store	8
Lip	Long-lasting color and shine in a compact portable palette	8
Eye	The 18-hour long-wear eye shadow that won't smudge, run, or fade	7
Lip	Won't smudge, run, or transfer…so you can eat, drink and kiss without having to reapply	7
Color	Small light beams can sense the difference between matte and glossy and detect the finest nuances in color	6
Segment 2—Interested in extra features, *Techie*		
Eye	Available at your neighborhood drug store	29
Color	With an additional car charger.so you can charge and go, always there when you need it	20
Eye	A bonus leather case keeps it clean and protects it from heat	20
Eye	Find it in the beauty section of department stores nationwide	20
Eye	With a bonus brush for error-proof application	19
Color	Comes with a rechargeable battery.for a seemingly endless life	18
Color	Sits safely in a cushioned case.keeps it out of harm's way	18
Color	Find it in the electronics section of department stores nationwide	18
Eye	Buy it directly from the manufacturer's online Website	18
Eye	Enhanced with alpha hydroxy and fruit acids…improve skin's texture	18
Color	Additional removable memory chip stores shades and hues.so you can keep all the colors you create	16
Color	So reliable, it comes with an extended 10-year warranty	16
Eye	Pearl extract adds brilliance and luminosity to lids	15
Eye	With a unique blend of oil-free moisturizers…so your eyelids feel silky smooth	15
Lip	Satin feel and finish…for creamy, moist-looking lips	15
Lip	Buy it directly from the manufacturer's online Website	11
Lip	Moisturizing, long-lasting lip color…perfect for any skin tone	11

Table 3: cont….

Segment 3—What can be accomplished with the technology		
Color	Small light beams can sense the difference between matte and glossy and detect the finest nuances in color	31
Lip	For the woman who wants to make a personalized statement reflective of her identity	31
Color	Utilizes over a billion hues of color.discover the perfect shade with ease	30
Lip	Perfectly pouted lips…lets the classic woman release the dramatic side	30
Color	The latest technology to detect and match colors beyond the range of human vision	26
Lip	Won't smudge, run, or transfer…so you can eat, drink and kiss without having to reapply	25
Color	Connects to your computer, PDA and a variety of other devices.the perfect color companion	24
Eyes	Shimmery color kisses lids…adds a bit of glamour to your look	23
Color	Available in your neighborhood technology and electronics dealer	22
Color	For the technology-savvy individual looking for a new toy	22
Eyes	A dazzling collection of six perfectly harmonized eye shadows.bring out the best in your eyes	21
Eyes	A racy palette with six dramatic shades in one slim compact	21
Lip	Juicy lips with brilliant color…create a runway look in one stroke	20
Color	Accurate color differentiation.match colors as precisely as possible	18
Lip	For those who like a natural look with beautiful color…enhance your look in just a few seconds	18
Eyes	The perfect way to contour, highlight and define eyes…adds the final touch to your appearance	17
Lip	Potent color and luxurious feel…for the daring and seductive woman	17
Eyes	Rich, long-wearing and crease-free…your eyes deserve the best	16
Color	High resolution and accuracy.easily and precisely distinguish between shades of colors	15
Color	Find it in the electronics section of department stores nationwide	15
Eyes	Perfect to wear day and night…perfect for any occasion	15
Lip	Purchase with the aid of a personal sales representative in the comfort of your home	15
Lip	For the woman who's always on the run…easy to use wherever you are	15
Color	Purchase with the aid of a personal sales representative in the comfort of your home	14
Color	A pocket-sized color detector.to help identify the right colors	14

Table 3: cont….

Eyes	Made with hypoallergenic ingredients for contact wearers and sensitive eyes…dermatologist and ophthalmologist approved	14
Eyes	The 18-hour long-wear eye shadow that won't smudge, run, or fade	14

Note. The elements are sorted in descending order by utility value.

It is important to keep in mind that the real goal of segmentation is to discover a structure of the mind, namely different segments of people's mind-sets, rather than identifying any individual as belonging to one of these three segments. That is, we are using the segment to identify these mental archetypes in the cosmetic area. Thus, the segmentation approach proposed in Mind Genomics® represents a crossover between conventional segmentation done by marketers and archetypes-based thinking done by psychoanalysts (Wertime, 2003). The segmentation is an operationally straightforward, defined method for uncovering these archetypes or locations of mind-sets. Segmentation involves the way the mind organizes the information, rather than the way people divide into groups. Such an approach using conjoint analysis and segmentation as a method for identifying locations of ideas in a mind-space rather than people appears to have been first promoted in the automobile sales business by Moore and Moskowitz (2002).

Step 6: Selecting Idealets to Recombine into New Products and Running the Fourth (Recombinant) Study

Recombining creates newer and better concepts, not necessarily for a single product, but even perhaps for a new-to-the-world product. Our three studies on skin color sense, eye shadow and lip gloss allow the developer to create such a recombinant idea. We begin with a basic positioning statement—namely a product that enables the user to understand her skin tone, the appropriate eye shadow and the appropriate lip cream. We don't necessarily know what this product will be, as there are no rules for a new-to-the-world product. However, we can present winning *idealets* from the three studies, as shown in Table **3**. These *idealets* win among different segments.

The second stage of the project comprises a new study, this time with 36 elements, selected from winning ideas in the first phase, but selected from the three initial (*i.e.* basic) studies. Let's now put these *idealets* into an underlying structure or architecture as we did for the basic study and test combinations of these *idealets* as we did before. We simply introduce the new product idea by the basic positioning statement, not forcing the participant into any pre-defined mental framework. We then present different, systematically varied combinations of these elements. The combinations are mixed and matched. The positioning statement ensures that the participant knows that the product idea deals with a personal electronic make-up palette.

In the actual study a total of 6,000 *new* participants were invited by e-mail, with 260 individuals participating. Time frames dictated the completion of the study within 72 hours, which decreased the response rate to 4.3% of the invitees. Each new participant evaluated a unique set of 48 combinations, much in the way that the previous participants had evaluated a unique set. The 36 elements came from the three different studies so that the orientation and rating question had to be couched in general terms. Keep in mind that the participants had no idea that the elements were really abstracted from previous studies; all they knew was that they were evaluating a presumably "reasonable" idea based upon the introductory positioning. The participants were again segmented into three groups to identify different mind-set positions.

The partial results for the study appear in Table **4**, which shows the performance of the winning elements for the three segments developed from the new data. We tried to use the same names as were used for the first part of the study, although there were some differences, especially in Segment 1. In the first set of studies Segment 1 comprised individuals interested in short messaging and basic benefits, whereas in the fourth (spliced elements) study this segment comprised individuals interested in bottom line performance. Such study to study variation in segmentation should not surprise, given the differences in positioning and elements.

Step 7: Identifying Interactions Among Pairs of Ideas to Prevent Poor Combinations

Before creating a new combination of ideas by splicing together components it is important to determine whether the combinations "work" together or whether they do not. Some combinations make intuitive sense, some combinations do not. These combinations may be identified ahead of time and specified as pairwise restrictions. However, there are many combinations that just do not seem to "work" together, even though there is no reason, *a priori*, to assume that they would fail to work. It may be that to participants in the study the combinations are counterintuitive, or clash with each other, even though one would never have guessed.

Table 4: Results from the Second Phase (Study #4), With Elements Selected from Three Different Products and With the Concept Positioned Simply As an "Electronic Palette".

	Total	**Concept response segment**		
		Perform	**Techie**	**Usage**
	100%	42%	26%	32%
	260	109	68	83
Additive constant	36	39	12	51
Segment 1—Bottom line oriented—super performance				
Utilizes over a billion hues of color.discovers the perfect shade with ease	2	12	0	–9
The 18-hour long-wear eye and lip colors that won't smudge, run, or fade	7	11	11	–2
The latest technology.detects and matches colors beyond the range of human vision	5	9	12	–5
Segment 2—Interested in extra features, *Techie*				
Additional removable memory chip stores shades and hues.so you can keep all the colors you create	5	7	24	–13
Buy it directly from the manufacturer's online Website	–1	–11	17	–4
A bonus leather case keeps it clean and protects it from heat	5	7	16	–7
Sits safely in a cushioned case.keeps it out of harm's way	0	–1	15	–13
Find it in the beauty section of department stores nationwide	3	2	15	–6
Available at your neighborhood drug store	5	2	14	1
With an additional car charger.so you can charge and go, always there when you need it	1	–1	14	–8
Available in your neighborhood technology and electronics dealer	–8	–15	14	–17
So reliable, it comes with an extended 10-year warranty	7	8	12	2
The latest technology.detects and matches colors beyond the range of human vision	5	9	12	–5
Comes with a rechargeable battery.for a seemingly endless life	2	4	11	–9
The 18-hour long-wear eye and lip colors that won't smudge, run, or fade	7	11	11	–2

Table 4: cont….

Find it in the electronics section of department stores nationwide	–7	–11	11	–16
Accurate color differentiation.match colors as precisely as possible	0	0	11	–8
Small light beams sense the difference between matte and glossy and detect the finest nuances in color	3	4	10	–5
Moisturizing, long-lasting eye and lip color…helps you find the perfect shades for any skin tone	7	7	10	4
Perfectly pouted lips and bright, striking eyes.lets the classic woman in you release your dramatic side	5	7	9	1
Long-lasting color and shine in a compact portable palette	6	5	8	5
A dazzling collection of eye shadow and lip cream colors.brings out the best in your lips and eyes	3	2	8	–2
Segment 3—What can be accomplished with the technology				
Easy to apply, easy to remove…beauty in a matter of seconds	8	7	4	11
Mix, layer, lighten or intensify…achieve the perfect shade	5	7	–6	10
Brilliant color.create a runway look in one stroke	4	4	–2	9
A pocket-sized make-up palette with a built-in color detector.helps you identify the right colors	6	4	5	8
Make-up you can wear all day without having to reapply	6	8	1	8

Note. Only “winning” elements are shown for each segment.

Fortunately, the permuted, main-effect experimental designs used in these studies allow the discovery of significant interactions, both of positive and negative natures, respectively. The approach is quite simple, uses the principles of statistics and follows these steps to quickly reveal which combinations do better than expected and which combinations do worse:

Data Preparation. Line up all of the raw data, comprising rows of 36 columns (one per concept element) and a 37th column corresponding to the rating on the 9-point scale. With 48 concepts per participant and with 260 participants there are 12,480 rows.

“Interest” Measure. Create the 38th column corresponding to interest, where interest takes on the value 100 if the rating is 7–9 to denote interest, or takes on the value 0 if the rating is 1–6 to denote lack of interest.

Create All Pairs of Elements from Each of the Two Silos. There are six silos, A–F, so there are 15 pairs of silos [(6 × 5)/2 = 15]. For each pair of silos there are 36 pairs of elements (*e.g.,* A1…A6 crossed with B1…B6 generates 36 combinations). Therefore there are 15 × 36 or 480 pairs of elements.

Identify What Pairwise Interactions Co-Vary With Interest. Compute the Pearson correlation (or other measure of association) between each element pair and the interest value. There are 480 of these correlations, one per element pair.

Rank Order These 480 Interactions and Consider Only Those With Strongly Significant Positive Correlations (> 0.025) and Negative Correlations (< –0.025). These are the combinations that synergize so that the combination of elements does far better than chance, or that suppress so that the combination does far worse than chance. These seeming low correlations are, in fact, quite significant when one realizes they are computed using 12,000+ observations.

Use the Negative Correlations as Constraints. When it comes time to identify winning combinations, make sure that no poor-scoring combinations enter. These would combinations of elements that might perform well alone, but do not do well together.

Table **5** shows these synergistic and suppressive combinations for the total panel. Only the most significant pairs are shown.

Step 8: Synthesis of New Ideas Using a Recombinant Optimizer

A key benefit of genomics-based thinking is that ideas can be recombined into newer and possibly better combinations. The splicing of ideas exists already in the basic design of the research, where the elements are treated as individual pieces and recombined by the computer program during the course of the interview. Once the utility values of these individual ideas are identified, it becomes possible to further recombine the winning ideas into yet newer concepts by mixing together winning ideas. The analysis of interactions discussed above will warn whether the combinations that look promising on the basis of individual elements have a negative utility when combined. Judgment works as well, indeed in parallel with statistics, when deciding what combinations of optimal elements make business sense.

Table 5: Pairs of Concept Elements That Either Suppress Each Other (First Set of Elements With Negative Correlations) or Synergize With Each Other (Second Set of Elements With Positive Correlations).

Suppressive Combinations				**Synergistic Combinations**			
Pair	**Pearson *R***	**First Element**	**Second Element**	**Pair**	**Pearson *R***	**First Element**	**Second Element**
B3–F3	–0.037	When you want to make a personalized statement reflective of your identity	Purchase with the aid of a personal sales representative in the comfort of your home	A1–D5	0.025	Long-lasting color and shine in a compact portable palette	Utilizes over a billion hues of color.discovers the perfect shade with ease
E4–F3	–0.037	Additional removable memory chip stores shades and hues.so you can keep all the colors you create	Purchase with the aid of a personal sales representative in the comfort of your home	A2–F4	0.026	Moisturizing, long-lasting eye and lip color…helps you find the perfect shades for any skin tone	Find it in the beauty section of department stores nationwide
A5–F3	–0.033	A racy eye shadow/lip cream palette with a variety of dramatic shades in one slim electronic compact	Purchase with the aid of a personal sales representative in the comfort of your home	A3–F1	0.026	A pocket-sized make-up palette with a built-in color detector.helps you identify the right colors	Available at your neighborhood drug store
D5–F3	–0.032	Utilizes over a billion hues of color.discovers the perfect shade with ease	Purchase with the aid of a personal sales representative in the comfort of your home	B5–D2	0.026	For those who like a natural look with beautiful color.enhance your look in just a few seconds	The 18-hour long-wear eye and lip colors that won't smudge, run, or fade
B5–F3	–0.030	For those who like a natural look with beautiful color.enhance your look in just a few seconds	Purchase with the aid of a personal sales representative in the comfort of your home	B5–E4	0.026	For those who like a natural look with beautiful color.enhance your look in just a few seconds	Additional removable memory chip stores shades and hues.so you can keep all the colors you create
C2–F3	–0.030	Mix, layer, lighten, or intensify…achieve the perfect shade	Purchase with the aid of a personal sales representative in the comfort of your home	B5–C6	0.027	For those who like a natural look with beautiful color.enhance your look in just a few seconds	Easy to apply, easy to remove… beauty in a matter of seconds
C3–F3	–0.029	Brilliant color.create a	Purchase with the aid of a personal sales representative in the comfort of your home	C6–F4	0.027	Easy to apply, easy to	Find it in the beauty

Table 5: cont….

		runway look in one stroke				remove…beauty in a matter of seconds	section of department stores nationwide
E3–F3	–0.029	Sits safely in a cushioned case.keeps it out of harm's way	Purchase with the aid of a personal sales representative in the comfort of your home	E4–F4	0.030	Additional removable memory chip stores shades and hues…so you can keep all the colors you create	Find it in the beauty section of department stores nationwide
A4–F3	–0.028	A dazzling collection of eye shadow and lip cream colors.brings out the best in your lips and eyes	Purchase with the aid of a personal sales representative in the comfort of your home	C2–F4	0.033	Mix, layer, lighten or intensify…achieve the perfect shade	Find it in the beauty section of department stores nationwide
B1–F3	–0.028	When the technology-savvy side of you is looking for a new toy	Purchase with the aid of a personal sales representative in the comfort of your home				
C6–F3	–0.027	Easy to apply, easy to remove…beauty in a matter of seconds	Purchase with the aid of a personal sales representative in the comfort of your home				
D1–F3	–0.027	Small light beams sense the difference between matte and glossy and detect the finest nuances in color	Purchase with the aid of a personal sales representative in the comfort of your home				
E2–F3	–0.027		Purchase with the aid of a personal sales representative in the comfort of your home				
A1–F3	–0.026	Long-lasting color and shine in a compact portable palette	Purchase with the aid of a personal sales representative in the comfort of your home				
D3–F3	–0.026	Satin feel and finish.for creamy lips and fabulous eyes	Purchase with the aid of a personal sales representative in the comfort of your home				
E1–F3	–0.026	With an additional car charger…so you can charge and go, always there when you need it	Purchase with the aid of a personal sales representative in the comfort of your home				

Table 5: cont….

C5–F3	–0.025	Wear each shade alone, or combine shades to turn everyday eyes and lips into irresistible eyes and lips	Purchase with the aid of a personal sales representative in the comfort of your home				
E6–F3	–0.025	A bonus leather case keeps it clean and protects it from heat	Purchase with the aid of a personal sales representative in the comfort of your home				

Note. The suppressive pairs should not appear together in the same concept.

Structure of the Concept. The concept comprises only three elements, not six elements. The reason for this constraint comes from the fact that in the actual evaluations participants evaluated concepts comprising a minimum of three and a maximum of four elements. Even though there were six silos we look only at the best set of elements with three silos, with only one element from each silo. The other three silos are missing. For this optimization the three silos selected for the optimization were C (mode and ease of use), D (features and benefits) and E (additional features, merchandising).

How to Create the Combination. The combination is created without paying attention to constraints that might be imposed from knowledge of how the combinations of elements performed. However, we can check from Table **5** whether the optimal combination comprises elements that do not work together. There are no combinations that would correlate negatively with interest, suggesting that the three pairs of elements in the optimized concept are compatible with each other. The pairs are C6–D2, C6–E5 and D2–E5, which we construct by knowing that the optimal combination comprises C6, D2 and E5.

How Well Does The Concept Perform? The sum of the utilities is 57, coming from an additive constant of 36 and three utilities each of value 7. The modest value for total panel should not surprise since the segmentation suggests that the population is not homogeneous. Rather there are groups with diverging interests, so what appears to one group of participants may be very unappealing to another. For this particular combination of C6, D2 and E5, the concept scores as follows:

1. Total panel: 57
2. Segment 1 (techies): 65

3. Segment 2 (performance): 39
4. Segment 3 (usage): 62

Analyzing Subgroups. The same type of analysis may be done for any subgroup or set of subgroups, forcing in any triple of silos, or even allowing the computer to pick the silos based upon the attempt simply to optimize interest. For example, when we focus on Segment 2 (add-ons) we can generate another combination, C6, D6 and E4. The acceptance goes up from 39 to 52, at the expense of acceptance by the other segments.

DISCUSSION

It is well known that self-assessments of importance are often tremendously flawed, as shown in a comprehensive monograph presenting the results of more than 200 published articles on self-assessment (Dunning, Heath and Suls, 2004). These flaws, which could reduce the validity of utility values for individual ideas, manifest themselves empirically. For example, the utility values of well-known brands are much lower in concepts than the utility values of statements about product features (Moskowitz *et al.* 2005c). Even though brands are assumed to be very important, brand names, *i.e.* surrogates from brands, show relatively low utility values ranging from –10 to +5, for literally dozens of well-known brands in studies performed both in the USA and Europe (Germany, France, UK) and among both teens and adults. The disconnect between brand names as they perform in concepts (*i.e.* vignettes or mini advertisements) and the commonly held conceptions of brand names when assessed alone in the absence of anything else makes one wonder about how valid are stand-alone assessments of ideas. *In any event, conjoint analysis is not the science, but simply the best method "today" for obtaining the data on which the science of Mind Genomics® is based.*

The Foundational Points of View Underlying the Science of Mind Genomics®.

We can summarize the foundations of this newly proposed science of Mind Genomics® in the following points that provide both the specifics of the method perspective from the sciences that comprise the foundation.

The organizing principle of stimulus–response, taken from experimental psychology, enables the researcher to understand the "mind," based upon the patterns of reactions to "cognitively rich" stimuli.

People do not know necessarily what is important to them but can react intuitively to ideas. When these ideas comprise systematically varied vignettes (combinations of elements or *idealets*), then statistical analysis using regression reveals determine which specific concept elements or *idealets* "drive" the consumer responses.

Deep insight comes from exploring responses assigned an intuitive level, rather than from responses assigned after rational consideration.

A strong understanding of what is important to consumers comes from presenting them with a large set of such systematically varied combinations and getting them to respond at an intuitive or "gut" level, not at a considered intellectualized level. This strategy of research more naturally approximates what happens in the external world.

A series of studies, investigating the different aspects of a product, service, or life situation teaches far more than does any single study.

Any domain (*e.g.,* food preferences, states of anxiety, financial services) can be better dealt with through a series of such experimentally designed combinations, rather than one single study alone. Only through such studies can we investigate the granularity of life, the specifics, in the way that life is experienced. Thus, when it comes to the Mind Genomics® of food we might wish to have a dozen to five dozen such studies, each of which deals with a different food or different eating condition, with each study dealing at the granular level of detail. This view that one can gain a broader view of the consumer mind-set comes directly out of the science of genomics, where the researcher obtains a sense of how a gene expresses itself through multiple tests, not just one test.

A common structure across many RDE studies (so-called meta-study design) generates more knowledge and deeper knowledge because the structure allows for comparability and databasing.

The different tests (*i.e.* for different foods) are best laid out in a single common structure, with the specific *idealets* in each test particularized to the product being studied. However, the nature of each test element is specified by a template, so that the researcher can discover immediately how the same exact element or similar type of element performs across studies.

Respondents should be encouraged to participate in studies about topics that are interesting and relevant.

The studies should be set up so that an individual is invited to participate in the general project (*e.g.,* food cravings, healthy food products, insurance, anxiety states). Only when the individual expresses interest in a particular topic does it make sense to guide the individual into the specific study. In this way, the researcher ensures that the respondents who participate have selected themselves as being interested. Only later do they actually go into a specific study, through a second selection process, to participate in the topic-specific study. Such an approach means that the data generated in Mind Genomics® projects will come from the relevant individuals. As the elements are cognitive rich, representing meaningful stimuli in a person's everyday life, it makes a great deal of sense to work with people to whom the topic is relevant. Such an approach differs dramatically from studies where people are recruited because they fit into a specific demographic group.

Analysis of responses to the systematically varied test stimuli generates a model showing the impact of each test element for each respondent. With all elements considered, these impact values or utilities provide a sense of the "mind" of the respondent.

These individual utility values, in the aggregate, show the "mind-set" of the respondent to the category and constitute a "footprint" of the respondent's mind. By changing the rating attribute, we learn about how "instructions to the mind" change the respondent's point of view and judgment criteria. By changing the test framework (*e.g.,* type of food) we learn about how the same mind responds to a variety of similar types of stimuli (*i.e.* similar messages across foods). By working with many individuals in the population with the same rating scales and the same test stimuli, we identify the nature of different mind-sets in the population (mental genotypes). The mind-sets may be specific to a single product or may transcend a set of related products so that the mind-sets become an organizing principle for the larger product category. The mind-sets become the structure for understanding, not the people.

New Ideas can be Generated by "Mashing Up" Idealets into New Combinations.

This Lockean approach to concepts holds that the science of Mind Genomics® is both normative, revealing what exists and prescriptive, suggesting through recombination what could be. Such prescriptive approaches are very important for advancing the science of consumer research, especially in the commercial world, where development can be done using knowledge about the consumer mind-set.

Applications of Mind Genomics®.

Our goal in founding this new science is to better understand the value structure of the individual's mind using high-level consumer research tools. The mega studies comprising related studies in a product category provide an overview to the way consumers make trade-offs among options in the category. Looking across different studies provides insight into the distribution of mind-sets or mental genotypes worldwide.

Mind Genomics® has another objective: practical application of the knowledge and insights to create better products and services. Our suggested new science, therefore, stands atop two platforms—knowledge about people's judgment criteria when it comes to "ecologically meaningful" stimuli such as products, as well as direct application of the results in a business framework such as communication and product development.

We have applied the approach of Mind Genomics® ™ to areas as diverse as food craveability, beverages, insurance, anxiety-provoking social issues and shopping. Our next goals are to take the approach and apply to areas as diverse as the morals/ethics, political policy and financial issues. The goal of this early stage research is to show proof of the concept, develop a database, show how the results can be used and build this newly developing science from the "ground up".

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

This chapter is an adaptation of Moskowitz, H., Gofman, A., Beckley, J. and Ashman, H. (2006). Founding a new science: Mind Genomics®. *Journal of Sensory Studies, 21*(3), 266–307.

DISCLOSURE

Part of the information included in this chapter has been previously published under the title Founding a New Science: Mind Genomics® in *Journal of Sensory Studies* (2006), 21, 266-307.

REFERENCES

Aarts, P., Paulus, K., Beckley, J., & Moskowitz, H. R. (2002). Food craveability and business implications: The 2002 EuroCrave™ database. Paper presented at the ESOMAR Congress, Barcelona, September, 2002.

Ashman, H. & Beckley, J. H. (2004). Getting food away from home…it's convenient! Institute of Food Technologists Annual Meeting and Food Exposition, Las Vegas, NV, July.

Ashman, H., Beckley, J., Adams, J., & Mascuch, T. (2002). Teens *versus* adults: The 2001 Teen Crave It Study. Institute of Food Technologists, Anaheim, CA, July.

Ashman, H., Rabino, S., Minkus-McKenna, D., & Moskowitz, H. R. (2003). The shopper's mind: What communications are needed to create a "Destination Shopping" experience? Proceedings of the ESOMAR Conference "Retailing/Category Management—Linking consumer insights to in-store implementation". Dublin, Ireland, October.

Ashman, H., Teich, I., & Moskowitz, H. R. (2004). Migrating consumer research to public policy: Beyond attitudinal surveys to conjoint measurement for sensitive and personal issues. ESOMAR Conference: Public Sector Research, Berlin, Germany, May.

Beckley, J. H. & Ashman, H. (January/February, 2004). Understanding consumers' fears and anxieties about obesity and their impact on products. *Cereal Foods World, 49*(1), 39–41.

Beckley, J. H., Gillette, M., & Marketo, C. (2002). Changes in consumer mindsets since 9/11—The Crave It!® Foundational Study (sponsored by McCormick & Company). Food & Beverage Innovation 2002—Trend-focused strategies for product success. IIRUSA, Chicago, IL, November.

Beckley, J. H., Ashman, H., Maier, A., & Moskowitz, H. R. (2004a). What features drive rated burger craveability at the concept level? *Journal of Sensory Studies, 19*, 27–48.

Beckley, J. H., Ashman, H., Maier, A., & Moskowitz, H. R. (2004b). Hispanic and non-hispanic responses to concepts for four foods *Journal of Sensory Studies, 19*, 459–485.

Beckley, J. H., Moskowitz, H. R., & Minkus-McKenna, D. (2004c). Compelling innovation between industry and academia. Yale School of Management JMR Competition on Academic-Practitioner Collaborative Research, New Haven, CT, December.

Box, G. E. P., Hunter, J., & Hunter, L. S. (1978). *Statistics for experimenters.* New York: Wiley and Sons, Inc.

Collins F.S., Green E.D., Guttmacher A.E., Guyer M.S.(2003). US National Human Genome Research Institute. A vision for the future of genomics research. *Nature,* 422: 835-47.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education and the workplace. *Psychological Science in the Public Interest, 5*(3), 69-106.

Gofman, A. & Moskowitz, H. (2010). Isomorphic permuted experimental designs and their application in conjoint analysis. *Journal of Sensory Studies, 25*(1), 127–145.

Green P. E. & Rao V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research, 8*, 355–363.

Green, P. E. & Srinivasan, S. (1990). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5*, 103–123.

Himmelstein, J., Ashman, H., Moskowitz, H. R., Minkus-McKenna, D., & Rabino, S. (2004). The algebra of the customer mind-decoding drivers of customer choice for over-the-counter drugs and supplements in the healthcare arena. Global Healthcare 4 Conference, Paris, France, February.

Hirsch, J. B. & Zawel, S. A. (2002). Healthy You! Nutrition ingredients that drive consumer demand in healthy foods. Institute of Food Technologists, Anaheim, CA, July.

Hughson, A., Ashman, H., de la Huerga, V., & Moskowitz, H. R. (2004). Mind-sets of the wine consumer. *Journal of Sensory Studies, 19*, 85–106.

Krieger, B., Ashman, H., & Mascuch, T. C. (2002). Innovative ideas in the dairy business from the consumer mind: Update from the Crave It!® study. Presented at IQPC, Chicago, IL, October.

Luce, R. D. & Tukey, J. (1964). Conjoint analysis: A new form of fundamental measurement. *Journal of Mathematical Psychology, 1*, 1–36.

Luckow, T., Aarts, P., & Moskowitz, H. R. (2003). A comparison of purchasing habits and sensory preferences for cola consumers across France, Germany, the United Kingdom and the United States. *Journal of Food Technology, 1*, 84–96.

Luckow, T., Moskowitz, H. R., Beckley, J., Hirsch, J., & Genchi, S. (2005). The four segments of yogurt consumers; preferences and mind-sets. *Journal of Food Products Marketing, 11*, 1–22.

Minkus-McKenna, D., Ashman, H., & Moskowitz, H. R. (2004). Diabetes products: What health care marketers need to know to improve effectiveness of the shopping experience. *International Journal of Medical Marketing, 10*, 119–128.

Moore, P. & Moskowitz, H. R. (2002). Towards a new paradigm: Using advanced Internet research technologies to explore and optimize the preferred offerings of CarMax Inc., a “New Age” auto retailer. Proceedings of the Automotive Conference, ESOMAR, Lausanne, March.

Moskowitz, H. R. (2001). Customer driven development: A paradigm for sustainable innovation in “Cyberspace” using quantitative research methods. *Australasian Journal of Market Research, 9*, 15–26.

Moskowitz, H. R. & Ashman, H. (2003). The mind of the consumer shopper: Creating a database to formalize and facilitate the acquisition and use of insights. Proceedings of the ESOMAR Conference "Excellence in Consumer Insights". Madrid, Spain, March.

Moskowitz, H. R., Ashman, H., Gillette, M., & Adams, J. (April, 2002a). Moving closer to the consumer though understanding the mind: The Crave IT!® Study and the specific case of chocolate. *Ingrediente Alimentari,* 13–21 (in Italian).

Moskowitz, H. R., Beckley, J. H., & Adams, J. (November/December, 2002b). What makes people crave fast food? *Nutrition Today, 37*(6), 237-242.

Moskowitz, H. R., Itty, B., Ewald, J., & Beckley, J. H. (2004). The calculus of consumer privacy. ESOMAR Conference–Integrating marketing research in business: From managing data to generating decisions. Lisbon, Portugal, September.

Moskowitz, H. R., Porretta, S., & Silcher, M. (2005c). *Concept research in food product design and development*. Ames, IA: Blackwell Professional.

Moskowitz, H. R., Silcher, M., Beckley, J. H., Minkus-McKenna, D., & Mascuch, T. (2005d). Sensory benefits, emotions and usage patterns for olives: Using internet-based conjoint analysis and segmentation to understand patterns of response. *Food Quality and Preference, 16*, 369–382.

Poskanzer, D. & Ashman, H. (2003). Functional beverages—What is the difference between performance beverages and healthy beverages? Institute of Food Technologists Annual Meeting and Food Exposition, Chicago, IL, July.

Suppes, P. & Zinnes, J. L. (1963). Basic measurement theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology*, Vol 1 (pp. 1–76). New York: Wiley and Sons, Inc.

Systat. (1997). Systat, the system for statistics. *User's manual*. Evanston, IL: Systat Division of SPSS.

Van Ommen, B. & Stierum, R. (October, 2002). Nutrigenomics: Exploiting systems in biology in the nutrition and health arena. *Current Opinions: Biotechnology, 13*(5), B517–B521.

Watkins, S. M. & German, J. B. (2002). Metabolomics and biochemical profiling in drug discovery and development. *Current Opinions: Molecular Therapy, 4*, 224–228.

Wertime, K. (2003). *Building brands & believers: How to connect with consumers using achetypes*. New York: Wiley and Sons, Inc.

Wittink, D. R., Vriens, M., & Burhenne, W. (1994). Commercial use of conjoint analysis in Europe: Results and critical reflections. *International Journal of Research in Marketing, 11*, 41–52.

PART VII

EXPERIMENTATION AND CONSUMER-DRIVEN INNOVATION

CHAPTER 21

Consumer-Driven Innovation Management

Arcot Desai Narasimhalu[1,*] and Shekhar Mitra[2]

[1]Institute of Innovation and Entrepreneurship, Singapore Management University, Singapore and [2]Proctor and Gamble, Cincinnati, OH, USA

Abstract: The evolution of human society leads to increased affluence and prosperity of certain populations, sometimes at the expense of well-established markets. Market leaders in products and services tend to be so focused on their current customer base that they are caught off guard with the changes in markets created by the evolution. These changes often go unnoticed until it is too late. The change in customer base often requires the repositioning of products and services through innovations, which address new and emerging markets. Some of these changes could potentially result in tectonic market shifts that force innovation managers to involve current and future customer bases in order to help understand the opportunities that can lead to innovation. The nature of these innovations could span the range from addressing the mundane needs of developing countries to meeting the wishful aspirations of mature markets. Firms are often at a loss on when and how to use customer-inspired insights in the goal to create new innovations. Innovation management takes on a new art form that engages customers, allowing them to reveal their unmet needs. Such a fuzzy front-end process demands new engagement styles and structures that are less obvious to those who use traditional tools such as surveys and focus group research. This chapter identifies the challenges faced by firms in responding to a less-traversed approach toward using customers to identify innovation opportunities, and suggests methods to manage such challenges.

Keywords: Innovation, consumer-insights, fuzzy front-end.

BACKGROUND

Have you wondered what Boeing, BMW, Coors, Electronic Arts, IKEA, Lego, and Staples have in common? All of them have started engaging their customers in one form or another to identify their innovation-led growth opportunities.

Many companies achieve market leadership from one or more initial innovations. They even catapult themselves into the Fortune 500 or equivalent star lists. And then something sad happens. Even the best of innovations, however well

***Address correspondence to Arcot Desai Narasimhalu:** Institute of Innovation and Entrepreneurship, Singapore Management University, Singapore; Tel: 65- 6828 0914; Email: Desai@smu.edu.sg

Alex Gofman and Howard R. Moskowitz (Eds)

protected, attract competition. The market leaders set their sights on the competition, and a new, hitherto unsuspected, competitor creeps into their market. This focus, often entirely on the company's acknowledged, major competition is normally the beginning of the end of a company's ability to sustain its market leadership. The only sustainable competitive advantage any company can enjoy is its ability to stay focused on its customers and increase its customer base by creating a string of innovative products and services.

Companies traditionally had either their marketing departments or their planning or strategy departments worry about the next set of products and innovations, which would help them retain their market leadership. This approach usually worked when the innovations were incremental. If a company was producing black and white television sets, it would make sense for it to consider producing color television to satisfy the demands of their customers, who were craving enhanced experiences.

There are often disruptions in either technology or the markets, which require a totally different approach to creating innovations. For example, if a firm creating shampoos left the planning of its next innovations to its marketing department, it would most likely end up creating next generation shampoos with additional functional or emotional appeal for developed economies. However, it was those who saw the product through customers' eyes who discovered that shampoo in a sachet as a new innovation was very powerful to expand their customer base, and hence their revenue base and profits (Prahalad, 2005). Understanding customer requirements remains an extremely important step in a company's desire to create innovations that will succeed.

Understanding and identifying customer requirements has been a holy grail for the product and service designers and managers. The task of identifying customer requirements can be approached from multiple perspectives. A robust approach would be to use Maslow's hierarchy of needs as a framework to identify customer needs (Maslow, 1943). Another approach would be to understand cultural diversities to identify innovations. A third approach could be to use the differences between needs and wants in order to identify customer requirements. A fourth approach might be to understand how innovations have evolved in the past and use this knowledge to identify customer requirements for the next set of innovations. Finally, one could use open innovation (Chesbrough, 2003) as a mechanism to identify customer needs. No matter which of these paths pursued to identify customer needs, it is very important that the customers themselves are integrated into the identification process. It is those products that can be truly called customer-led innovations.

The concept of co-creation was introduced in 2000 (Prahalad and Ramaswamy, 2000). Consumer-inspired innovation is but one of the many forms of co-creation. The objective is to get the product developers and the lead users to be actively engaged in identifying innovation opportunities. In 2009, Promise Corporation worked with LSE Enterprise to sharpen the definition of co-creation into "co-creation is an active, creative and social process, based on collaboration between producers and user…initiated by the firm to generate value for customers" (LSE Enterprise and Promise Corporation, 2009).

In this chapter, we discuss co-creation or consumer-inspired innovation from each of the perspectives listed in the previous paragraph. The use of Maslow's hierarchy of needs will be discussed next. After that, we will discuss co-creation using cultural diversity, and then the differences between the needs and the wants and the methods of involving consumers in creating innovations in each of the categories. Innovation evolution paths will be introduced next, and then we will discuss the use of abduction in consumer-inspired innovations. Finally, we will touch on the use of open innovation in consumer-inspired innovation and present some examples from Proctor and Gamble (P&G). We conclude with a summary of our discussions.

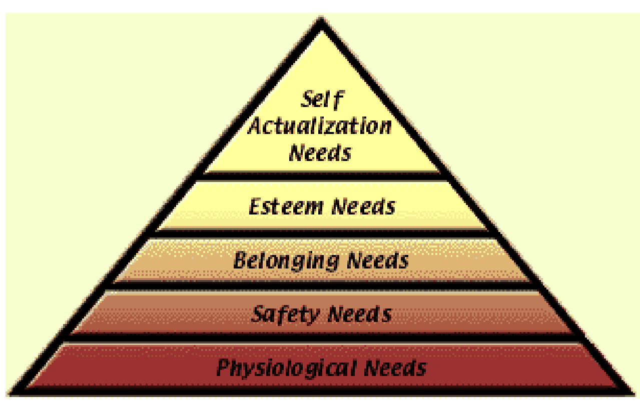


Figure 1: Maslow's hierarchy of needs.

MASLOW'S HIERARCHY OF NEEDS

Abraham Maslow identified five different levels of human needs starting from physiological needs and ending up with needs for self-actualization as presented in Fig. **1**. Each of the five categories of needs have been further studied and their respective subcategories have been defined, as shown by Table **1**.

Table 1: Some Subcategories of Maslow's Hierarchy of Needs.

Level	Need category	Need subcategories
5	Self-actualization	Creativity, morality, objectivity, open-mindedness, problem solving, spontaneity
4	Esteem	Achievement of vision, confidence, respect by others, respect for others, self-esteem
3	Love/Belonging	Family, friendship, intimacy, member of a community
2	Safety	Emotional, family, personal, physical, possessions, professional, social
1	Physiological	Air, food, water, shelter

A preferred approach to consumer-inspired innovation first ascertains the level in the hierarchy of needs toward which a product or services is currently positioned. Then, one involves consumers to identify specific needs for innovations at the higher levels of the hierarchy.

The question arises regarding what might be the best means of engaging the customers in identifying such needs. Popular approaches such as focus group research or surveys represent well-tested instruments for revealing a person's explicitly understood needs. Oftentimes, a need remains latent until it is revealed, and thus discovered. It is in such instances that one needs to employ other observation-based techniques.

When an innovation addresses a physiological need such as shelter, it is time to involve consumers to discuss their needs at the safety level, *i.e.* securing their shelter. After the safety needs have been met, then the company can address innovations at the next level—Belonging.

For example, a product innovation at the "belonging" level may be a new type of community hall, a meeting space where the residents of a community can come together and which is accompanied by service innovations such as providing novel entertainment to the community. Once the needs at the belonging level are delivered, innovations at the "esteem" level can be addressed. An example might be service innovations for leadership training or presentation skills. At every level consumers need to be involved to determine their specific needs. Once the needs at the esteem level are addressed, the needs at the self-actualization level can be explored. An example could be creativity-related training.

It is important to note in this example that the consumers are involved in every step of the innovation process, and at every level. The involvement of consumers can happen both proactively or reactively. For example, someone who has built a shelter may proactively design a mechanism for securing the shelter in consultation with consumers. Alternatively, a security-related innovation might be in response to an unfortunate event such as a robbery. Although securing a shelter is a need, it might lie dormant or latent until a robbery takes place. "What if" analyses or scenario planning are useful tools to identify latent needs of a customer base, especially in cases where one does not wish to wait for unusual, rare, or unfortunate events to happen as a spur to innovation.

CULTURAL DIVERSITY OF CUSTOMERS

The previous section discussed how the needs shift from one level of Maslow's hierarchy to the next and the importance of getting consumers involved. This section discusses how culture has an impact on the innovations.

In some cities of the world (based on a 1984 Tokyo experience), the houses were never locked. The neighbors were trustworthy, and there was very little robbery or thievery. Thus, securing one's house was interpreted as lack of trust of the neighbors. Contrast this with a house in a very insecure neighborhood frequented by robbers. No one would take offense in producing an innovation for securing a house in such a neighborhood. This is an example of how culture might influence the need for and the realization of innovations.

There are many dimensions to understanding cultural diversity, beginning with linguistic differences all the way to practices imposed by a religion. Take biometrics, for example. A fingerprint-based biometric identification system will face huge entry barriers in Islamic countries as well as in Japan, but for very different reasons. In Islamic countries, putting a finger on a scanner that was used by others might be considered "haram" or dirty. In contrast, fingerprinting was used in Japan at that time to identify criminals and foreigners, and hence it was not considered to be the "right innovation" for ordinary usage.

Sometimes, innovations are dictated by cultural or religious factors. For example, the Casio Islamic Prayer Digital Watch CPW-310, as presented in Fig. **2**, was invented to address the needs of the Islamic communities in identifying the direction of Qibla (direction that should be faced when a Muslim prays) from anywhere in the world. In addition, this watch also carried the Hijra calendar and alarms for prayer times. It is clear that such a watch could not have been designed without the involvement of the respective consumer base.

Figure 2: Casio Islamic prayer digital watch CPW-310-1 VDS.

DIFFERENTIATING BETWEEN NEEDS AND WANTS

Innovations can be created to address both the needs and the wants of consumers. "Needs" are "must have" requirements; hence, any innovation addressing needs are easily accepted by the target markets. In contrast, “wants” are “good to have” requirements; hence, any innovation addressing wants is much harder sell. Typically, start-ups tend to be the drivers of innovations that address the needs, whereas incumbents are the ones whose innovations address the wants.

Needs are things that human society must have for its existence. Examples are:

- Nutritious food
- Shelter (dwelling)
- Clothing
- Footwear
- Transportation

Wants are things that are good to have if one can afford it. Examples are:

- Designer clothing
- Toys
- Chocolates
- Video games
- Jewelry

Some innovations are wants that are, in truth, simply enhancements of needs. Table **2** outlines such instances using the five needs listed above.

Table 2: Innovations That are Wants Built Upon Needs.

Needs	**Wants**
Nutritious food	Gourmet food
Basic housing	Gated resort property
Clothing	Fur coats and haute couture
Footwear	Branded shoes
Transportation	Sports cars

Conversely, over time some wants can evolve into needs. Take as an example the automobile-reliant transportation in cities without adequate public transportation. In such cities, an automobile becomes a "must have" object.

Some wants tend to become needs as a society becomes more affluent. For example, the rural population in some countries could not afford footwear and hence walked barefoot. However, in countries that are more developed, footwear is a need and not merely a want.

The important take-away here is that wants and needs are contextual. They must be considered with reference to the geographies under consideration. Hence, it is important that companies involve consumers from the relevant geographies when they co-create innovations for specific geographies.

INNOVATION EVOLUTION PATHS

The innovation cube as defined in Narasimhalu (2005) provides a framework to represent the attributes of successful innovations and has evolved to incorporate innovation rules as a part of the innovation engine (Narasimhalu, 2007).

Fig. **3** presents the innovation cube. The innovation cube is defined by three dimensions—innovation drivers, innovation triggers, and innovation enablers.

Innovation drivers are either pain or pleasure. Pain suffered by a community of consumers or enhanced experiences (pleasure) sought by another community are drivers of successful innovations. Pain roughly corresponds to needs and pleasure roughly corresponds to wants. The deeper the pain and the more widespread it is, the greater the value created and market size. The same is true of the pleasure.

Innovation triggers are market shifts and technology shifts. Although a pain or pleasure is identified, the markets and the technologies must be ready. Otherwise, they remain as future innovations and do not have immediate promise. Such innovations succeed when the markets and technologies are ready.

Some innovations fail despite satisfying the innovation drivers and triggers. This is mainly due to their inability to scale to fit in the market or when the price is not right. The price and the speed of scaling were identified as innovation enablers.

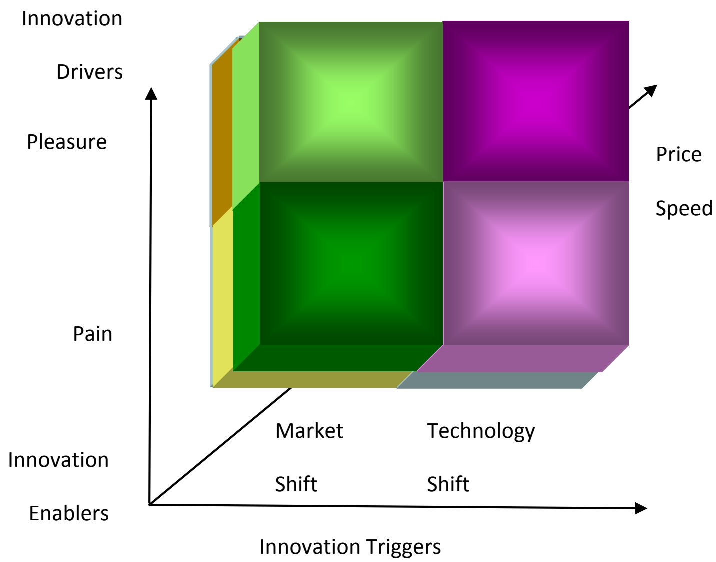


Figure 3: The innovation cube framework.

The innovation cube was used to derive innovation rules. The sample innovation rule presented in Fig. **4**, in a schematic format, shows the evolution of several product lines such as computers, fax machines, and copiers.

For example, computers were created for defense purposes and then moved on to provide enterprise-level solutions, department-level solutions, and desktop solutions, before ending up as laptop computers and PDAs or smart phones.

Every innovation rule has two or more stages or levels linked by arrows. When an innovation lies at a given level, e.g, the enterprise level in rule 1, market readiness and the technology readiness will trigger innovations at a division level. The rule would read: “If the markets and technologies are ready for the transition from the enterprise to the division level, then create the division-focused innovation. Otherwise, either create the technology required for the innovation, or if the technology is available then wait for the markets to be ready”.

One way to determine when markets are ready involves consumers. When an innovation is at the enterprise stage and the technology for transitioning it to the divisional stage is ready, then consumer inputs can reveal whether or not consumers have a need or a want for a division-level innovation. When the consumers long for a division-level solution then it is time to create such an innovation. Then, co-creation managers can work with consumers to apply innovation rules and identify new innovation opportunities.

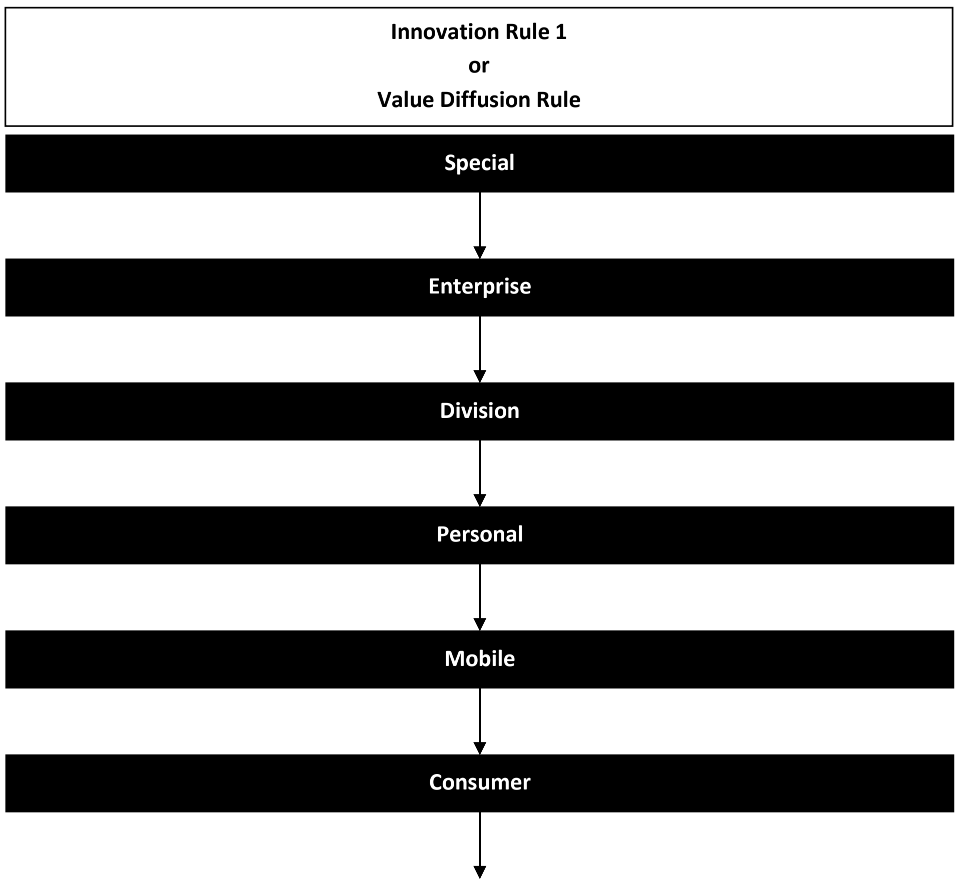


Figure 4: An example of an innovation rule.

USING ABDUCTION FOR CONSUMER-INSPIRED INNOVATION

Whereas deduction allows deriving "A" as a consequence of "B," and induction allows inferring the association between "A" and "B" from multiple occurrences of "A" and "B," abduction allows the inference of "A" as an explanation of "B". Abductive reasoning analyzes a set of seemingly unrelated facts in order to create the type of hypothesis sometimes referred to as a hunch.

When companies have a "hunch" that an innovation would be relevant to a target market, they could actively engage the target customers in order to fine-tune the manifestation or prototype of the innovation. For example, if a furniture design and manufacturing company were to plan products for single females, then it is likely to have a hunch on what type of product innovations might make sense for this target market. It would be prudent of this company to engage a collection of its unmarried female customers to help design the furniture for them.

The selection of the target group of single females will again have to be sensitive to the context of the target market. For example, whereas there might be some common requirements across all single females from different geographies, religions, and cultures, there could very well be a special/customized requirement for every subset of single females the company targets as its customers.

OPEN INNOVATION—ITS IMPACT ON CONSUMER-INSPIRED INNOVATION

Open innovation can take place at multiple levels as described in Narasimhalu (2008) and presented in Fig. **5**. Consumer-inspired innovations could be co-created from any of the levels of this model. For example, using consumers at large, consumer-inspired innovation can happen at the highest level in the open innovation model. In contrast, consumer-inspired innovation at the enterprise level will operate the level 4 of the open innovation model.

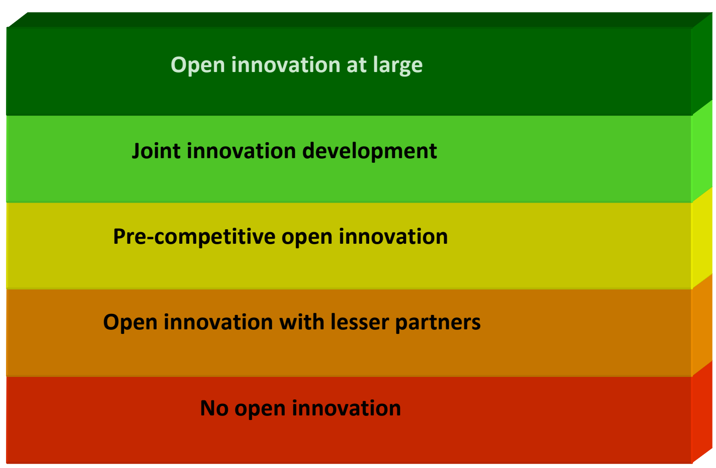


Figure 5: Multiple levels of open innovation.

Crowd-sourcing is another example of an open innovation method (Howe, 2006). Crowd-sourcing is, in fact, the best example of co-creating innovation with the involvement of consumers.

Companies can decide which of the levels of the open innovation models presented in Fig. **5** is relevant to a consumer-inspired innovation.

A Framework for Consumer-Inspired Innovation

Discussions of Maslow's hierarchy of needs, innovation evolution paths, cultural diversity, needs and wants, open innovation, and abductive reasoning all can be sewn into a framework for an innovation process inspired by consumers. Fig. **6** presents such a framework.

Maslow's hierarchy of needs, innovation evolution path, or any other method could be used to derive candidate innovations. These candidates can then be processed through the innovation engine in order to verify the readiness of innovation triggers, *i.e.* markets and technologies required for creating the innovations. Innovations for which both the markets and technologies are not ready will be stopped at this stage.

The candidate innovation that passes the innovations triggers test will then be passed through the "wants–needs filter" in order to determine whether it is a want-based or a need-based innovation. Want-driven innovations will have a target market that is generally much smaller than the markets of need-driven innovations. Once again, it is important to remember that needs and wants are contextual. A need sometimes crosses over to become a want, and *vice versa*.

At this juncture, wants-driven innovations and needs-driven innovations take two independent but similar paths.

The innovations are first passed through a "culture adaptation engine" to ensure that the innovations meet the cultural sensitivities and that culture-related issues do not pose any adoption hurdles. New requirements might be added to satisfy culture-specific requirements.

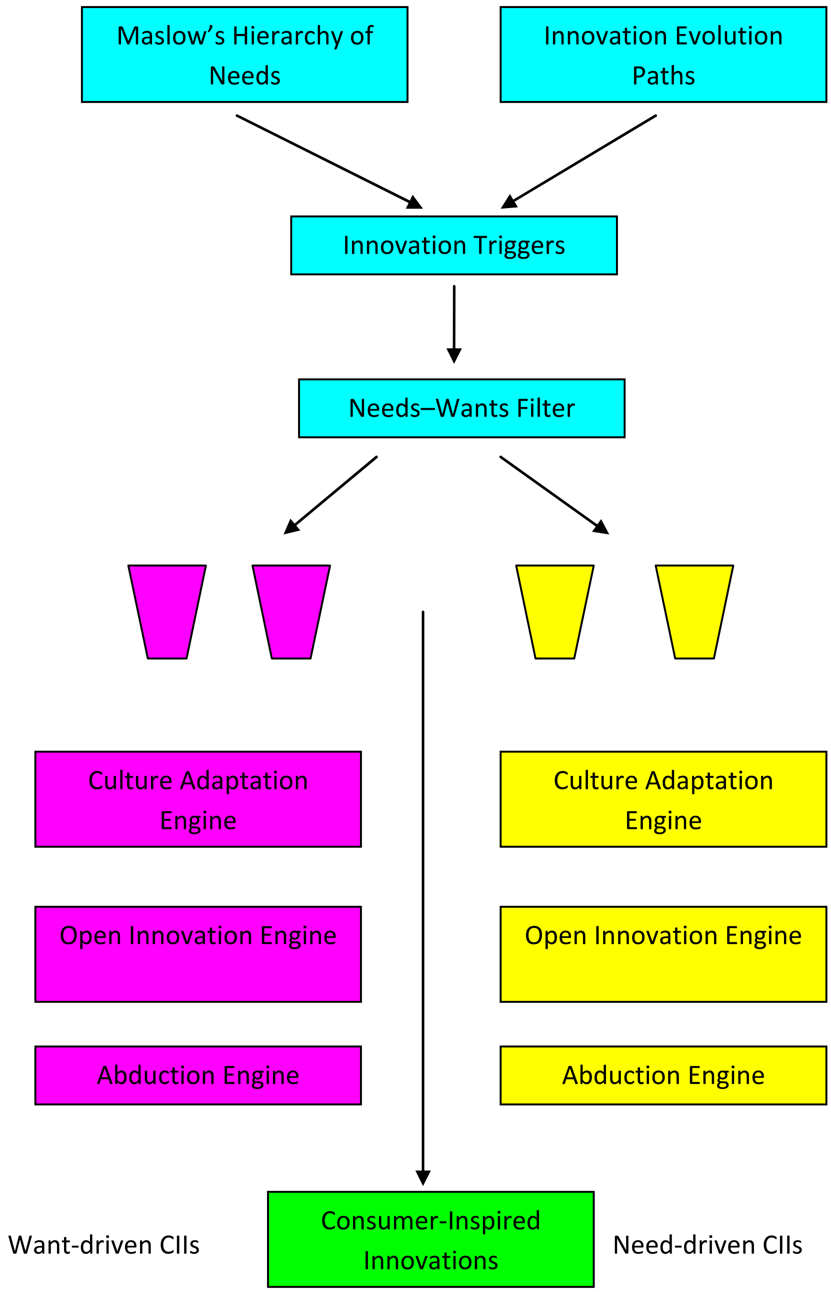


Figure 6: A framework for consumer-inspired innovation process.

The innovations that emerge from the culture adaptation engine are then passed through an open innovation engine. The open innovation engine can suggest the level of the open innovation model at which consumer/customer participation is recommended. The consumer-inspired innovation manager can then decide whether or not to accept the recommendation or even to alter the recommendation as deemed necessary.

The best recommended consumer engagement model is to observe consumers while they are in their natural environments where the innovation would be used. For example, if the innovation is related to a home television set, it is best then to observe consumers interact with the television in their homes. Such observations remove extraneous artifacts that might be introduced in a controlled environment.

In the absence of an opportunity to observe consumers in action, it might be useful to have someone talk to the consumer using either an audio or a video link. The interviewer ought to be trained how to ask open-ended questions that will emulate the in-person observation.

If all else fails, traditional survey instruments may be used. However, survey instruments are best avoided if at all possible given that most often respondents tend to provide answers that they think the surveyors would expect.

There is an increasing trend to build in some of these capabilities into customer relationship management (CRM) software. CRM software could very well be used to elicit customer inputs on a proposed innovation. CRM software can also be used to receive unsolicited suggestions for innovations from inspired consumers/customers.

All the inputs on a candidate innovation received from consumers can then be run through an abduction engine to formulate hypotheses on the consumers' responses; these hypotheses can be tested using social science research methods.

INNOVATION METHODS AND EXAMPLE FROM P&G

This section first presents the consumer-inspired innovation strategy used by P&G, and then introduces some examples of innovations introduced by P&G in different categories.

Developing Innovations Strategies

P&G is a leader in innovation in general and open innovation in particular. P&G plans its innovation projects with the end in mind, from the consumer perspective, typically working backward from a vision that is five years ahead. It then creates a storytelling initiative, which acts as the master plan. The initiatives are aimed at producing innovation outcomes with bigger, better, faster, and less costly value propositions.

P&G focuses on three types of innovations: (1) sustaining growth-oriented innovations, (2) disruptive market-oriented innovations, and (3) commercial innovations.

Sustaining growth-oriented innovations (SGIs) are developed to fill the gaps in a product line, to eliminate trade-offs, to offer new benefits, or to eliminate product negatives. Gillette Fusion Power is an example of an innovation in this category.

Disruptive market-focused innovations (DMIs) create entirely new sources of consumption that are likely to introduce new users by perhaps even cannibalizing current markets. The aim of this category of innovations is to make the "impossible" possible by creating innovations that, over time, evolve into stand-alone product categories. The pitfall against which one must guard, is to avoid getting caught up in the desire to perfect an innovation, without, however, understanding what the target markets might consider to be good enough. P&G carefully monitors such innovations to ensure that such products have low knowledge/assumption ratios. Some examples of DMIs are Pampers diapers, Swiffer Wet Jet, and Crest White Strips.

Innovations that fall under the commercial category are generally market innovations without any product or package changes. These are designed to help provide constant news, encouraging new consumers to try the product and turn into loyal customers who purchase the product again.

P&G's disciplined, scientific approach to innovation often includes a life-cycle assessment of a product. P&G helped pioneer this research tool, which looks at environmental factors such as carbon dioxide, energy and water consumption, and waste over the entire lifespan of a product from raw materials, to product manufacture and logistics, to consumer use, and to the final post-consumption disposal. This comprehensive approach helps P&G identify the biggest opportunities to improve the environmental impact of its products.

Profoundly Understanding Customers

P&G focuses its efforts on delighting what it calls the "sustainable mainstream consumer". This group of customers typically comprises 75% of reachable consumers in each of its key geographic regions.

P&G has made a substantial investment in understanding its customers' beliefs, habits, and what drives their purchasing decisions. P&G tracks newly emerging definitions of value in the mind of the customer to guide development. The value triad comprises three critical elements: performance, price, and sustainability profile. The customer generally will not and often does not sacrifice performance or price for environmental benefits, especially in tough economic times.

P&G has defined internal criteria for "sustainable innovation products". Innovation falling under the sustainable products label ought to:

- Reduce usage by more than 10% in resources such as energy, water, transportation, packaging, and product material without trading off benefits in other indicators.

- Be supported by good science that is substantiated by data and must be verified by their stringent claim approval systems.

P&G deploys the three-stage DID (define, invent, demonstrate) model to design, develop, and deliver its innovations. Table **3** captures the tasks carried out in the three stages for the sustaining growth-oriented and disruptive market-focused type of innovations.

Innovation Examples From P&G

NA Laundry Compaction Product

An example of sustainable product innovation is P&G's NA Laundry Compaction product from its fabric care division. The product innovators doubled the concentration of their liquid laundry formula. The result was a holistic product redesign, generating high-impact benefits across the entire product life cycle. The new product was delivered in smaller packaging because of increased concentration. The annual savings were clearly measureable: 15,000 metric tons of packaging material; 40,000 fewer truckloads in shipping; savings of 500 million liters of water; and reducing carbon dioxide emissions by 100,000 metric tons in addition to the tacit benefit of the consumers having to deal with smaller packages. This also resulted in less inventory space for P&G, their distributors, and dealers. The outcome was substantially more sales. The shelf space required at the retailers' shelves were also almost halved.

Table 3: Proctor and Gamble's Innovation Methodologies.

Stage and SGI	DMI
Define	
• Identify consumer targets and their desired customer experience (DCE) • Develop the corresponding consumer concept/idea • Validate the business attraction/opportunities • Identify key inventions needed to address in the invent stage	• Select new domain • Identify the new consumers for the disruptive market innovation • Define the tasks to be carried out
Invent	
• Create solutions • Resolve areas requiring invention • Create solutions to killer issues • Screen through technologies • Develop proof of concept • Identify killer issues for development & resolution at demonstrate stage	• Develop the new business model • Resolve two to three assumptions to knowledge transfer
Demonstrate	
• Define the winning link between consumer need, innovation concept, product design, and underlying technology • Resolve major issues related to integration and development	• Ensure that all components of the business model work profitably • Identify early adoption markets that will build to inflection point

Ariel Turn to 30°

Ariel Turn to 30° is an example of a commercial innovation. The brand "pleaded" with the UK washing machine users to turn their water heating choice from 43.5°C to 30°C. The campaign explained how 80% of the energy in a washing machine is used to heat the water, and that there would be little difference in turning down the heat. A key communication point was that the savings allow the consumers to watch around 1,500 episodes of their favorite soap opera, or boil enough water to make 2,600 cups of tea. This campaign resulted in 17% of the consumers turning their heating choice to 30°, up from 2%. The result prevented the creation of 58,000 metric tons of CO_2 emissions. In 2007, this number went up to 30%. This campaign positioned P&G as a responsible thought leader for sustainable products and practices.

Charmin MegaRoll

Many innovations that are visible to consumers are also inspired by them. When Charmin bath tissue consumers expressed a desire to change the roll less frequently, P&G created Charmin MegaRoll, which features four times as many sheets per roll than a regular roll of Charmin. Along with meeting consumer needs, the Charmin MegaRoll required fewer cardboard cores, first for use, and then, of course, for disposal. In addition, the space-efficient product allows more tissue to fit on a truck, saving on fuel consumption and CO_2 emissions associated with transportation. Fig. **7** presents the impact from the Charmin MegaRoll innovation.

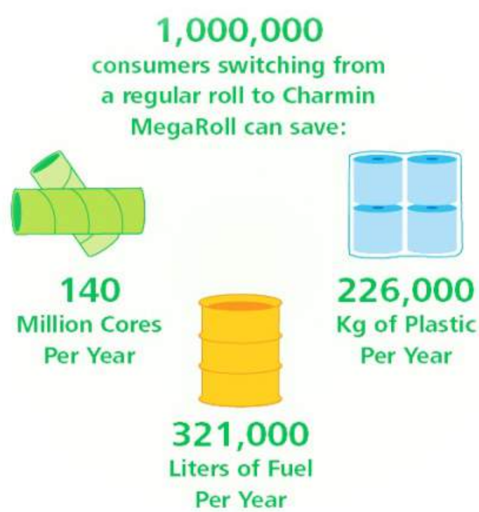


Figure 7: Impact of Charmin MegaRoll innovation.

Packaging Innovations

Consumer research on Prilosec indicated that getting the tablet from the blister pack was the number one consumer complaint. Novel design and technical innovation led to the development of an easier-to-access blister, which still metal safety requirements. This solution improved the consumer's experience while using Prilosec. By combining two blisters into a single blister, the brand saved more than 500,000 pounds of material annually.

Reduction in packaging can represent a step improvement in environmental sustainability. An example of such an innovation can be found in the Cover Girl cosmetics line. Consumer testing and market research of Cover Girl TruBlend facial foundation product showed that less packaging actually provided a better presentation to the consumer of the product on the shelf. It also led consumers to select the proper shades of their foundation with increased accuracy and satisfaction.

This research finding ended up eliminating the secondary packages, which surround the primary dispenser bottle leading to better product displays, easier shade selection, and more than 20% reduction in packaging.

Pampers

Pampers was invented in the 1960s with a breakthrough technology designed to deliver high-quality fit, softness in materials, and non-leaky product design. The product created a revolution in the market when introduced. However, as competition started entering this category, P&G realized that it needed to focus on the end user, *i.e.* baby and mother (in line with the "consumer is the boss" principle). P&G had to deliver a purpose-inspired innovation in addition to the technology. Furthermore, P&G had to communicate their product upgrades in terms of the difference and the improvement these upgrades offered to the baby's development. The core idea was to move away from comfortable, functional, and traditional promises to baby's happy, healthy development by connecting with mothers, *i.e.* digital marketing, mobile clinics for babies, *etc.*

In developing countries, the challenge was even more acute, especially regarding the need to change the habit from the use of cloth for diapers to the adoption of modern solutions. Mothers in these markets traditionally looked at Pampers as a "convenient" option especially for night usage, and they were least interested in using them regularly. Yet the moment P&G was able to move the product benefit focus in their communications from functional to baby care and development, the mind-set of the mothers changed, resulting in increased adoption for regular use. P&G drove consumer-inspired innovation and converted Pampers into a purpose-inspired and benefit-driven brand combining two key elements—caring for babies and mothers and delivering functionally superior benefits, respectively.

SUMMARY AND CONCLUSIONS

Consumer-inspired innovation is not a stand-alone concept. It takes its inspiration from works related to open innovation, crowd-sourcing, co-creation, and bottom of the pyramid. This chapter integrates the above concepts with Maslow's hierarchy of needs, innovation cube, and innovation rules to derive at a framework for open innovation. P&G's experience and commercial successes suggests that the framework has essential validity to guide ongoing innovation.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Chesbrough, H. W. (2003). Open innovation: The new imperative for creating and profiting from technology. Boston, MA: Harvard Business School Press.

Howe, J. (June, 2006). The rise of crowdsourcing. Wired. Retrieved March 17, 2007, from http://www.wired.com/wired/archive/14.06/crowds.html

LSE Enterprise & Promise Corporation (2009). Co-creation: New pathways to value. http://personal.lse.ac.uk/samsona/CoCreation_Report.pdf (retrieved June 14, 2010).

Maslow, A. H. (1943). A theory of human motivation. Psychological Review, 50, 394–395.

Narasimhalu, A. D. (2005, June). Innovation cube: Triggers, drivers and enablers for successful innovations. ISPIM Conference, Porto, Portugal, June 19-22, 2005.

Narasimhalu, A. D. (2007, June). *Innovation engine*, *ISPIM Conference*, Warsaw, Poland, June 17-20, 2007.

Narasimhalu, A. D. (2008, June). A maturity model for innovation management. ISPIM Conference, Tours, France, June 14-19, 2008.

Prahalad, C. K. (2005). The fortune at the bottom of the pyramid – Eradicating poverty through profits. Saddle River, NJ: Wharton School Publishing.

Prahalad, C. K. & Ramaswamy, V. K. (2000). Co-opting customer competence. Harvard Business Review, 78(1), 79-87.

Prahalad, C. K. & Ramaswamy, V. (2004). The future of competition, co-creating unique value with customers. Boston, MA: HBS Press.

CHAPTER 22

Neuromarketing 2.0: How Rule Development Experimentation is Innovating Neuromarketing Research

Sokol Zace[1,*] and Howard R. Moskowitz[2]

[1]Contact Design, Boston, Massachusetts, USA and [2]Moskowitz Jacobs Inc., White Plains, New York, USA

Abstract: The advances in neurosciences during the 1990s (the "decade of the brain") made it possible for the marketer to use cutting-edge brain imaging equipment, such as functional magnetic resonance imaging and ultra-high-resolution electroencephalography, literally to "look into" the consumer's brain and perhaps understand what physiological phenomena might be occurring. The first decade of the second millennium is characterized by an increasing media hype surrounding neurosciences as they are applied in marketing research (so-called neuromarketing). As often happens during pivotal moments of innovation in science and technology (the dot-com bubble being a recent example), some pundits of the sector run wild with speculations about the potentialities of neuromarketing. The neuromarketing promise can be summarized as follows: "Brain imaging and biometric techniques are capable of predicting the consumer's behavior". Regarding the neuromarketing promise, there are only a handful of peer-reviewed scientific research papers, compared with the number of pop culture publications. Even companies founded with the specific purpose of carrying out neuromarketing research are reluctant to give scientific references in order to back up their claims. Of course, they own proprietary methodologies that cannot be disclosed, but one might suspect that there is a lot of unfounded speculation and wishful thinking too. Therefore, it became critical to develop a scientifically based, rigorous method capable of discerning reality from myth, and of discovering new possibilities in neuromarketing. Rule developing experimentation (RDE), combined with EEG and eye tracking (treated in another chapter) turned out to be that method and worldview. This chapter describes the birth of the new science of RDE-based neuromarketing, and what it has told us thus far.

Keywords: Brain imaging, EEG, eye tracking, neuromarketing, RDE.

INTRODUCTION

How Consumers Make Their Daily Decisions

Neuroscience shows that often economic decisions may conflict with the usual

***Address correspondence to Sokol Zace:** Contact Design, Boston, Massachusetts, USA; Tel: (617) 265-7125; Email: sokol@ergolineus.com

model of rational choice based on the optimization of the relationship between costs and benefits. Despite the best mathematics of economists, and the best theory, the mind takes over. And the mind is not necessarily rational, despite the economic theory that should hold.

The complex interplay of daily life, and the different controlling mechanisms, both automatic and rational, appear in Table **1**. Most of the daily decisions depend on the interaction between all quadrants of the table.

Table 1: Factors That Determine Consumers' Daily Decisions.

Serial, require attention, can be summoned at will, permit introspective access	**Controlled**	Rational assessment of the meaning. Example: "This food is too caloric".	Recognize and reflect on the emotions that the stimulus has aroused.
Parallel, don't require attention, outside of conscious control, precede rational awareness	**Automatic**	Create a mental image of the stimulus. Example: "I recognize that the object in front of me is a food".	Emotional interaction with the mental image. Example: "I like this food".
		Cognitive	**Emotional**
		Cognitive processes allow us to collect information on the environment, store, analyze, evaluate, convert, and use them to take action. They answer the question "true or false".	Emotional processes motivate the behavior of attraction or avoidance.

Note. Courtesy of Francesco Gallucci, 1to1lab, Milan, Italy.

The Neuromarketer's Toolbox

The two main instruments neuromarketers use to study brain, cognition, and emotions are:

- The electroencephalograph (EEG)
- Functional magnetic resonance imaging (fMRI)

Historically, the EEG represents the oldest and most researched, noninvasive methodology used for the study of cognitive and emotional processes in the brain. The human electroencephalograph was developed by Hans Berger around 1924. The first encephalograph was a very large, overwhelming machine, as it had to be because of the electronics of its time. The EEG machine used needle and ink to plot electric pulsations emitted by the brain. The modern encephalograph is considerably smaller, varying in size but as small as a Walkman-sized encoder (see Fig. **1**).

Figure 1: The ProComp Infiniti™ EEG encoder by thought technologies.

Our brain produces electrical impulses all the time. Each impulse is characterized by its amplitude and by its frequency. The amplitude is the power of the electrical impulse, measured in microvolts (mV). The frequency is the speed of change of the electrical impulse, measured in cycles per second (or Herz, Hz). The frequency determines the category of brain waves—beta, alpha, theta, and delta. The combination of these categories determines or underlies our state of consciousness at any given time (Wise, 1995).

The EEG machines record the brain wave combinations (or patterns) by measuring their amplitudes and frequencies. The measurements are carried out through noninvasive electrodes placed in standardized positions on the scalp. The measurements can be made up to 2,000 times per second, generating enormous amounts of data. In truth, the EEG provides an abundance of data, much of which must be edited down to a manageable mass.

Placing sensors in different areas of the scalp allows electrical brain wave signals (rhythms) to be detected. These signals carry data about the levels of activation of the brain when it is involved in a functional interaction with any kind of communication stimulus. However, the "coded" information has to be deciphered. That is, simply measuring the electrical responses does not immediately tell us what is happening. For that we need experiments—careful observations of test conditions and responses—to begin the process of decoding what is happening.

Toward Understanding the "Code"

The basic concept in the EEG data analysis is that different states of mind and consciousness are associated with changes in EEG oscillations (rhythms) emitted by the brain. It is therefore crucial to create a method by which to characterize the oscillations in the EEG response, using well-defined methods with scientifically validated analyses. One of these methods, indeed a very popular one, is known as spectral analysis, where the word "spectrum" refers to a set of oscillations of the EEG signal at different frequencies. The analysis is automatic and does not require any interpretation up front. In that respect, spectral analysis is an objective approach with which to begin the decoding.

The "Basic Building Blocks" of the Code—What Does the EEG Provide?

EEG signals have frequencies ranging from 0.5 or 2 cycles per second (0.5–2 Hz) up to 40 cycles per second (40 Hz). EEG rhythms are divided on the basis of frequency ranges. Each specific frequency range is identified by a Greek letter. A frequency range is then defined as the frequency band. The brain-wave bands usually appear combined together, but in certain states of consciousness they may appear one at a time, or with one band prevailing over the others (Fig. **2**).

Let us consider each of the main brain-wave bands (beta, alpha, theta, and delta) separately in order to understand their meanings, and the meanings of their combinations.

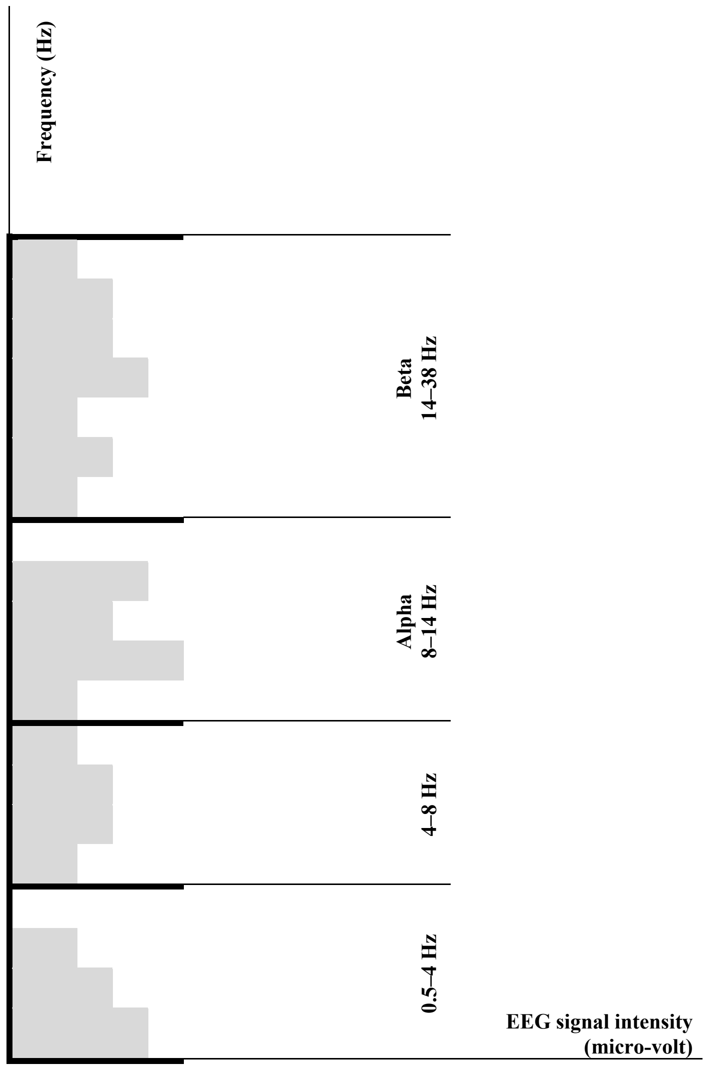


Figure 2: Main brain-wave bands. The gray areas represent an example of distribution of the EEG signal intensity along the frequency axis.

Beta Brain Waves

1. 14–38 Hz.

2. Cade and Coxhead, 1989) defines beta as “the normal waking rhythm of the brain associated with active thinking or active attention, focusing on the outside world or solving concrete problems. The strength of the signal is increased by anxiety and reduced by muscular activity”.

3. Beta indicates high arousal levels.

Alpha Brain Waves

1. 8–14 Hz.

2. Alpha brain waves are prominent during daydreaming, fantasizing, and visualization. Alpha waves also appear for brief moments before falling asleep, and during REM sleep.

3. They are associated with a relaxed, detached awareness and with a receptive mind.

4. From the point of view of interpretation, alpha provides the bridge between the conscious and subconscious mind.

5. The increase in the alpha wave's amplitude indicates greater availability to sensorial input.

Theta Brain Waves

1. 4–8 Hz.

2. Theta brain waves underlie activities of the parasympathetic nervous system. Most activities are carried out autonomously, without a conscious effort. They produce, or are associated with, theta brain waves. Autonomous fight-or-flight reactions are not associated with theta brain waves. They belong to the sympathetic nervous system.

3. From the point of view of interpretation, theta represents the subconscious, that part of our mind that forms a layer between the conscious and the unconscious. The theta frequencies are particularly intense during dreaming sleep and deep meditation. Hypnagogic imagery, those spontaneous and surprising images that appear out of the blue right before we fall asleep, is associated with high-intensity theta brain waves. These brain waves have been associated with access to unconscious material and creative inspiration. They indicate low arousal levels.

4. Last, but not least, the theta rhythm is associated with defocused attention, which is the cognitive process occurring when we scan the visual field waiting for something to appear, happen, or change, such as in an oddball videogame (Razumnikova, 2007). As we will see below, this property of the theta brain waves is of particular interest to neuromarketers.

Delta Brain Waves

1. 0.5–4 Hz.

2. Delta brain waves are primarily associated with deep sleep (without dream activity), pathological conditions, anesthesia, coma, or high emotional stress. Delta may also be present in a waking state in combination with other frequencies.

3. There are other brain-wave bands, such as mu, lambda, and gamma, which are left out of this presentation for the purposes of brevity.

EEG Indicators Used in Marketing Research

Typically, the following indicators are used in marketing research studies:

1. General (Defocused) Attention

A state of consciousness related to:

- Autonomous nervous system—actions performed automatically
- Parasympathetic activities—passive/receptive, as opposed to active/fight-or-flight
- Relaxed exploration
- All-inclusive (nonselective) perception
- Characterized by intense theta, and attenuated beta brain waves

This is the quiet explorer's state of mind—passive, relaxed, and alert at the same time—like the driver using peripheral vision to scan for a free parking lot, or the shopper walking through the aisles of shelves in a store. The person is ready for something that is yet to happen.

2. Focused Attention

A state of consciousness related to:

- Actions performed with a conscious effort
- Selective perception—one detail at a time
- Logic and reasoning
- Characterized by intense mid-band beta, and attenuated low-band beta brain waves

When active, conscious, logic, rational, and intellectual thinking takes place, the prefrontal cortex generates high-intensity mid-beta brain waves.

In terms of process, we can say at this stage that the searcher has already found the parking lot or a shampoo that might satisfy her or his needs. He or she is now evaluating if there is enough space to park there, or if the shampoo is compatible with the hair characteristics.

3. Memory Storage or Memory Encoding

A state of consciousness related to:

- Learning
- Updating of the cognitive schemes
- Middle-term and long-term memory
- Characterized by intense theta (Sederberg, Kahana, Howard, Donner, and Madsen, 2003), and attenuated mid-range beta brain waves

4. Evocative Activities

A state of consciousness related to:

- Retrieval of previous cognitive schemes, prototypes, and experiences
- Efforts of comparing the actual experience with the models stored in the long-term memory
- Characterized by intense delta and attenuated low-range beta brain waves

5. Decoding

A state of consciousness related to:

- Rumination: repetitively and passively focusing on the symptoms of distress and on its possible causes and consequences
- Characterized by intense upper-band beta, and attenuated low-range alpha brain waves

6. Anxiety

A state of consciousness related to:

- Reaction to excessive stress
- Difficulty in accomplishing a task
- Characterized by intense upper-band beta brain waves

7. Engagement

- The combined impact of indicators 1–4
- Measures the overall intensity of *positive* emotional and cognitive experiences during a determined time interval

8. Complexity

- The combined impact of indicators 5 and 6
- Measures the overall intensity of *negative* emotional and cognitive experiences during a determined time interval

9. Engagement S

This is an indicator that the one of the authors (SZ) found empirically. Subjects produced high-intensity theta and alpha waves at the same time, while watching stimuli that they subsequently described as likable, truthful, and reliable. They "bought into the idea", so to speak.

NEURAL HYPOTHESES ABOUT THE SHOPPING/SEARCHING PROCESS

In the world of neuromarketing there are a number of working hypotheses and rules of thumb that are accepted. Although these are not cast in stone, nor completely validated by studies in the literature, they make sense, and allow the neuromarketer to proceed with the analysis.

The subconscious and the conscious levels exchange a wide flow of signals in order to make sure everything fits. If a match of "feel good" and "make sense" is found, the parasympathetic nervous system generates a lot of theta brain waves, while the alpha brain waves are also prominent. A purchase decision is about to be made.

Of course, pricing of the product, and other analytical factors, have their own part in the decision-making process. The consumer may end up not buying the product, but engagement gives us an estimate of the purchase intent.

TECHNOLOGY AND DATA PRESENTATION

1to1lab uses "spider web" diagrams showing the value of the EEG indicators during each frame of a TV commercial, or for each shelf area looked at, during in-store marketing studies. Eye-tracking or head-mounted cameras are used "in synch" with the EEG, in order to identify the elements that are being looked at any time, while the brain is generating the measured brain wave patterns. The combination of the EEG diagrams with eye-tracking trajectory maps makes it possible to identify the high-impact scenes and the high-impact elements in a scene, as well as the low-impact, or negative-impact scenes.

The foregoing presents the technology, which allows one to discover what is happening. At the more proactive, engineering end, the marketer's goal is for the consumer to experience as many positive emotions and positive states of consciousness (attention, focus, memory storage, and evocation of pleasant memories), and experience these positive states for as long as possible. This criterion is used to compare different advertising messages, or different shelf set-ups, and reveal which specific solution generates these "positive responses". The assumption is that the solutions generating the more positive responses are likely to generate more sales.

FUNCTIONAL MAGNETIC RESONANCE IMAGING (FMRI)

The idea that certain areas of the brain are specialized in determined tasks dates back to a few centuries ago. A scientific study of the brain's compartmentalization became possible only in the second half of the 20th century. With the development of brain imaging technologies such as fMRI in the 1990s and through the study of brain damage, neuroscientists were capable of developing brain maps showing where motor, emotional, cognitive, and decisional activities took place (Bloom, Beal and Kupfer, 2002).

Functional magnetic resonance imaging is based on the following discoveries:

- Blood flow increases in active parts of the brain. The red blood cells alter the magnetic fields of those areas, making it possible to detect their location through an fMRI scanner. The fMRI technology does not require the injection of radioactive traces.

- The basic ideas underlying the application of fMRI in neuromarketing are similar to the ideas underlying EEG research: the activation of certain areas in the consumer's brain is supposed to give insight into the consumer's reactions to marketing stimuli.

GALVANIC SKIN RESPONSE (GSR) AND HEART RATE MEASUREMENTS

Galvanic skin response and heart rate measurements are gaining popularity among neuromarketers, especially in in-store marketing studies (Brat, 2010). The basic ideas underlying the application of these biometric studies in neuromarketing are similar to the ideas underlying EEG and fMRI research: certain patterns of GSR and heart rate oscillations are supposed to give insight into the consumer's reactions to marketing stimuli.

CRITICAL ISSUES

The late 2000s are characterized by an increasing consensus among neuromarketers on how to interpret the EEG indicators. One recurrent statement that can be found on neuromarketing firms' Websites (*e.g.,* NeuroFocus, 2010) and in articles about neuromarketing can be summarized as follows: "The EEG indicators of attention, memory storage, and engagement are capable of estimating the consumers' intention of purchase".

However, publications reporting efforts to verify the accuracy of EEG-based predictions through established, traditional research techniques (such as the research study reported by Sands Research, 2009) are still a rarity.

EEG is capable of showing with a certain degree of accuracy if the consumer viewing or listening to a marketing message is experiencing defocused attention followed by focus, anxiety, memory storage, and so on. However, the following questions should be addressed:

- Should we care?
- Are certain states of consciousness—or sequences of states of consciousness—better than others, from the advertising effectiveness perspective? Can we take it for granted that attention, focus, and memory storage are desired states of consciousness during a marketing message?
- Can EEG indicators predict purchase intent?

ENTER RULE DEVELOPMENT EXPERIMENTATION (RDE)

The RDE was developed by Moskowitz Jacobs Inc. (MJI), based upon advancing the methods of conjoint analysis. The beginning of RDE can be traced back to the early 1980s (Moskowitz and Gofman, 2007). The RDE research approach tests combinations of stimulus elements. These combinations are created according to an experimental design, to obtain data from consumers, generating models showing the relation between subjective responses and the independently varied stimuli. RDE finds use in both science and business because it systematizes the everyday world. The stimuli may consist of messages or actual physical variables.

RDE is built on these foundations:

- Experimental design, also known as design of experiments, or conjoint analysis
- Consumer ratings to the test stimulus
- Deliberate overwhelming of the participants' cognitive capacity, in order to *prevent them from guessing* what ratings would please the researcher
- Segmentation to identify different groups of people, or in actuality, different groups of ideas that naturally co-occur

The main steps of a typical RDE research study follow this sequence, as outlined by Moskowitz and Gofman (2007):

Step 1. Ideation stage: identify the problem, get ideas, edit them, and put the modified ideas into silos.

Step 2. Combine the elements into short test concepts (mix and match) and instruct customers to rate these different combinations.

Step 3. Use regression modeling to estimate the additive ratings caused by the contribution of each element.

Step 4. Using clustering techniques, segment or divide the respondents into groups based upon similar patterns of elements which drive responses. This is the mind-set segmentation.

Step 5. Write a set of rules of selection and combination of the elements, in order to generate the best-selling product, best-selling marketing message, or a combination of both. The R in RDE stands for "rule".

THE PILOT PROJECT—STEPS TOWARD THE EXPERIMENTUM CRUCIS (CRUCIAL TEST)

RDE provides a structure whereby the researcher can relate responses to well-defined physical stimuli. RDE, an intellectual descendent of psychophysics, searches for relations between variables just as psychophysics, a branch of psychology, searches for relations between well-defined physical stimuli and subjective responses.

Let's plug the EEG as a set of "responses" into the framework of RDE and attempt to uncover lawful relations. The short summary of the experiment comes from the group analysis by the four key experimenters. The goal of the summary is to show how RDE approached the problem.

Two caveats are in order:

- We are not going to prove or disprove the validity of neuromarketing. That job is impossible to do right now, at this early stage of development.
- We could have asked similar critical questions about any neuromarketing tool, such as fMRI, galvanic skin response, or other biometric techniques.

EEG as a response was selected as the key dependent variable. At the time of the experiment (2009), technology was already available to make EEG measurements relatively inexpensive. EEG was thus chosen for these three reasons:

- EEG devices are small, portable, not difficult to use, and relatively cheap (one unit typically costs a few thousand dollars).
- One stream of EEG raw signal, through mathematical transformations, generates a large number of indicators, as if they were measurements from "n" different devices. That stream guarantees richness of data, although it also requires establishing a filtering system to process only part of the data.

- EEG is the most widely used technology in neuromarketing for the time being and appears to be the most likely candidate for the foreseeable future.

We recruited 69 paid participants for one-on-one test sessions in a research facility in London, UK. The setup is shown in Fig. **3**. The early trial runs were filled with some artifacts. The cleaning drove the base size from 69 to 34, while at the same time revealing what to avoid when moving the EEG from a one-off study to a larger scale-up project.

Figure 3: Data acquisition session. The eye-tracking monitors display the stimulus, while the EEG signal is acquired by electrodes placed on top of the respondent's head (Cz Vertex position, according to the International 10–20 EEG system).

THE STIMULUS

The stimulus was the digital image of a yogurt package, with visual and textual elements that varied according to an experimental design. That is, the yogurt package was divided into different "areas". Each area was a variable, with alternative options. The experimental design dictated which particular options would join together to create a specific yogurt package. The experimental design thus generated many different yogurt packages. Fig. **4** shows the stimulus as the respondent saw that stimulus in color on a computer screen, totally assembled. In turn, Fig. **5** shows the same package, this time with the different *areas of interest* demarcated by rectangles. The visual and textual elements are automatically placed by StyleMap.net, the RDE tool developed by MJI.

Figure 4: An example of the test stimulus to be exposed to a respondent in an RDE study of packages.

The rationale behind an experimental design is the ability to relate the response (*i.e.* the EEG response) to the different elements. Rather than simply saying that one yogurt package generated a specific profile, and another yogurt package generated a different profile, experimental design permits the experimenter to show the precise relation between the features of the yogurt package and the EEG response.

SETTING UP THE YOGURT PACKAGES ACCORDING TO AN EXPERIMENTAL DESIGN—CONSIDERATIONS

We set up the experimental design using RDE principles. The combinations were set up so that there were five silos (areas of interest in Fig. **5**), and four alternative elements for each area of interest. Thus there were 20 elements altogether.

Figure 5: The template with the six areas of interest (AOI). Area of interest 6 is defined as any area not in areas 1–5.

The experimental design selected at most one element from each silo, but occasionally the design was set up so that no element from the silo was present. The experimental design was programmed so that all test packages comprised at least four elements (four out of the five silos or areas of interest were present). Most of the test packages comprised a complete set, one element from each silo.

The rationale for the incomplete design is simple. At the end of the interview, the RDE tool (StyleMap®) was programmed to create an equation relating the presence/absence of the 20 elements to the different EEG responses. In order to estimate the parameters of the equation, it was necessary to have some package combinations missing certain silos. By that stratagem, of having some "incomplete packages," it was possible to avoid the statistical problem of multicollinearity.

We are going to use ordinary least-square regression (OLS). The requirement of testing 20 elements, making them statistically independent, and having stimuli maximally missing one silo required a modified statistical design comprising 48 combinations.

RUNNING THE EXPERIMENT WITH RESPONDENTS

The respondents, recruited to participate for the study, began with an orientation. Respondents typically do not know what to do in these types of RDE studies, even though to the experimenter there seems to be no ambiguity. To ensure that all of the respondents knew precisely what to do, the experiment was set up so that the respondent began with an orientation screen, shown in Fig. **6**. The screen told the respondent exactly what to do.

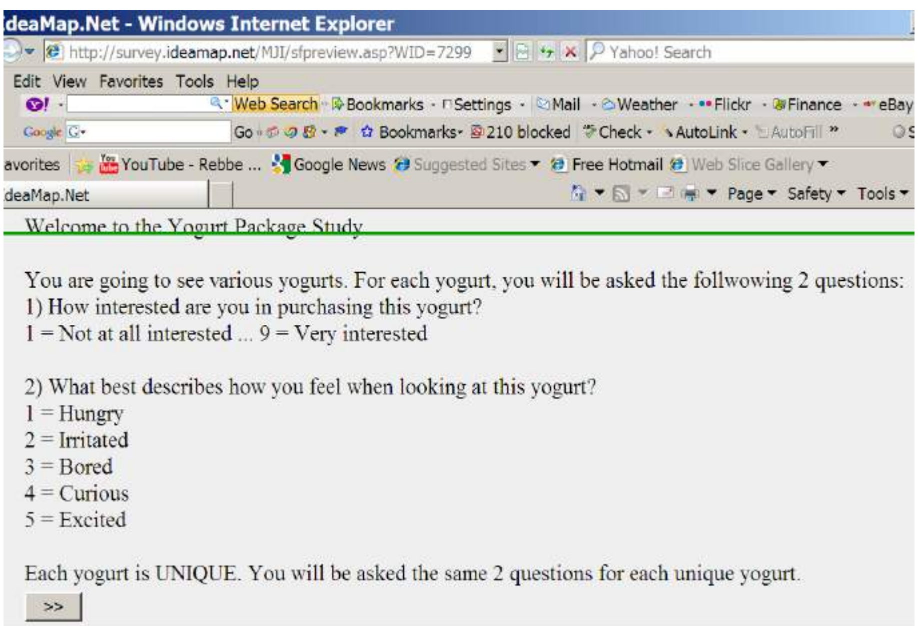


Figure 6: Orientation page for the yogurt study.

The experiment obtained three sets of measures: subjective ratings from the respondent (interest, emotion), EEG response, and eye movement, respectively.

For the subjective ratings, RDE used two rating questions. The first rating question attempts to get at an overall evaluation of the package.

1. How interested are you in purchasing this yogurt?

   1 = Not at all interested … 9 = Very interested

The second rating question requires the respondent to select ONE emotion that best described his feeling while looking at the yogurt package.

2. What best describes how you feel when looking at this yogurt?

   1 = Hungry

   2 = Irritated

   3 = Bored

   4 = Curious

   5 = Excited

For the EEG we used the eight-indicator setup, provided by 1to1 lab, of Milan, Italy (see "EEG indicators used in marketing research" earlier in this chapter)[1]. Although the different indicators are presumed to reflect specific functions, for the sake of scientific objectivity, we treated these leads as if they constituted different machines, whose function was to be determined through the experiment. The EEG information was sampled every half second, from 0.5 to 3.0 seconds, making this sampling synchronous with the eye-tracking information. At each half second, the EEG response from each of the eight leads was recorded.

For the eye tracking, we used the Tobii eye-tracking system, which showed which area of interest on the yogurt package was being looked at the same half-second intervals. That is, the eye tracker was synchronized with the EEG measurements.

CODING THE DATA—THE FIRST AND CRITICAL STEP TO UNDERSTAND WHAT WAS HAPPENING

The above-mentioned collection of data generated the following information, sufficient for an investigation into the interrelations among different classes of variables (package features, subjective responses, eye movement patterns, and EEG electric brain activity):

a. The elements

b. The ratings (interest, selection of emotion)

c. Eye movement for six sampling periods (0.5 through 3.0 seconds)

d. EEG responses from eight leads for the six sampling periods

With 34 respondents, each seeing 48 vignettes, rating each vignette once on attributes, but with six samples of eye tracking and EEG (once per half second), the experimental design generated $34 \times 48 \times 6 = 9{,}792$ records.

To understand the EEG results, we begin with the raw data. Our goal is to relate the different EEG responses to the ratings assigned by the respondents. The yogurt packages are merely devices by which we can elicit two classes of responses: subjective ratings and EEG potentials, respectively.

Our working data consisted of 9,792 cases. The data comprise results from 34 respondents, each of whom evaluated 48 different package designs. The respondents evaluated each of the 48 systematically varied packages once, assigning each an overall rating and a selection of the appropriate emotion.

For each test package we have six measures for each EEG lead. We have measures at 0.5–3.0 seconds, with a measure each half second. We can further average the six EEG measures per lead, to generate one single EEG measure per lead, corresponding to the average of the six time measures. This strategy reduces the 9,792 cases to 1,632 cases, or one-sixth of the data.

Each case, *i.e.* a specific experimentally designed package evaluated by a specific respondent, comprises the following information:

1. The respondent ID number.
2. The 20 elements, coded 1 when present, 0 when absent. For any case (specific respondent, specific package design) expect to see at most five 1's, but sometimes four 1's. Simply looking at the 1's and 0's tells us immediately what the package would look like.
3. The 9-point rating of interest.
4. The conversion of the interest rating to a binary (rating 1–6 map converted to 0, rating 7–9 converted to 100).
5. The specific emotion selected.
6. The eight EEG leads (each averaged over six times).
7. The location of the pupil of the eye from the eye-tracking data.

ANALYZING THE EEG INDICATORS THROUGH RESPONSE–RESPONSE (R–R) ANALYSIS

After preparing the data, usually the most time-consuming part of the process, we are ready to compute the correlations between the magnitude of electrical impulses from each EEG lead and each of the subjective ratings.

We have eight EEG leads (across the top of Table **2**).

We have two acceptance ratings, the original 9-point rating and the 0/100 binary conversion (1–6 → 0; 7–8 → 100). The latter approach comes from the world of consumer research, which looks at responses as yes/no.

We have the five emotion/feeling values, which take on the value 0 when the emotion/feeling was not selected and 100 when the emotion/feeling was selected. And of course, we compute the correlation based on the 1,632 cases.

The results in Table **2** disappoint. They show no strong correlations at all. That is, when we look at the entire data set and compute the correlations on the raw data (after averaging across the six time periods), we find no correlation at all—at least not yet.

Table 2: Pearson Correlation Coefficients Between the Magnitude of the EEG Response From Eight Leads, and the Subjective Ratings of Acceptance, the Binary Transformation of Acceptance to Interest, and the Binary Response to the Selected Emotion.

	Lead 1	Lead 2	Lead 3	Lead 4	Lead 5	Lead 6	Lead 7	Lead 8
Acceptance								
Rating	–0.01	0.01	–0.03	0.03	–0.01	0.01	0.01	0.04
Interest	–0.03	0.02	–0.03	0.04	0.00	0.01	0.00	0.06
Emotion/Feeling								
Hungry	–0.01	0.02	–0.04	0.02	0.09	0.02	–0.01	0.06
Irritated	–0.03	–0.03	–0.02	–0.05	0.04	0.02	–0.06	–0.01
Bored	–0.01	0.00	0.01	–0.01	–0.03	–0.02	0.00	0.06
Curious	0.04	0.02	0.03	0.02	–0.04	–0.02	0.05	–0.10
Excited	0.01	–0.02	–0.01	0.02	–0.06	0.03	0.01	–0.02

Note. 0 when not selected for the case, 100 when selected for the case.

Analyzing raw data obscures the potential patterns between subjective responses of interest and emotion *versus* the electrical response from the eight EEG leads. Let's move beyond the disappointing results from the raw data into an analysis that is potentially more productive and ultimately more insightful. This analysis works with the 20 elements. We will reduce our data set from thousands of observations to 20 observations, and do our analysis on the 20 observations or cases.

1. The basic unit of data will be the element from the package. There are 20 such elements. Each element will generate a profile of interest level (a rating), frequency of emotion (a rating choice), and EEG response from the eight EEG leads. We will end up with 20 cases, each case having numbers from ratings, and from the EEG.

2. The numbers will be obtained by regression analysis. There will be several regression analyses. The first regression analysis will relate the presence/absence of the 20 elements of the yogurt package to the rating of interest. The second to sixth regression analysis will relate

the presence/absence of the 20 elements of the yogurt package to the presence/absence of each of the five emotions. The seventh to fourteenth regression analyses will relate the presence/absence of the each of the 20 elements of the yogurt package to the EEG response from one of the eight leads. In the seventh to fourteenth regression analysis we work with the average of the EEG across the six sampling times (0.5–3.0 seconds).

3. Combining the information in Steps 1 and 2 above, we have 20 "cases," one per element. In each case we have the interest measure, the distribution of responses to the five emotions/feelings, and the EEG response of each lead, ***factoring out*** *the separate contributions of respondent and time*. We end up with 14 measures (one interest, five emotions/feelings, eight EEG leads), each with the same 20 cases.

4. We now compute a simple Pearson correlation on appropriate pairs of these measures, using our data from the 20 cases. We have gotten rid of a lot of the raw data for each measure, boiling down the measure to its essence, attributable to a specific element in the package vignette.

5. Our correlations are much higher now, giving us a sense of the strength of a linear relation between pairs of variables (Table **3**).

6. By so doing we see, for example, that the leads co-vary with different subjective attributes. In Table **3** we have shaded and bolded those cells with the absolute correlation 0.60 or higher. There are nine of those correlations across 48 cells. When we reduce the criterion to 0.50 or above for the absolute correlation, we add an additional nine cells, for a total of 18, or more than one out of three.

7. We find a majority of the EEG leads (five of eight) co-varying with the emotion of curious, and none co-varying with the emotion/feeling of "hungry".

8. Surprisingly, we find no strong correlation of any of the EEG leads to estimated interest from the interest model.

Table 3: Correlations Among Outputs of the Eight EEG Leads (Columns), and the Results From the Model for Interest, and for the Five Emotions/Feelings (Rows).

EEG Lead	Presumed "Meaning from Other Studies"	Interest	Curious	Bored	Irritated	Excited	Hungry
Lead 6	Anxiety	–0.50	**–0.80**	**0.73**	**0.60**	–0.29	–0.57
Lead 8	Complexity	–0.50	**–0.75**	**0.69**	0.57	–0.36	–0.49
Lead 3	Memory storage	0.50	**0.65**	–0.59	–0.57	0.45	0.40
Lead 2	Focus	–0.56	–0.26	0.42	0.49	**–0.69**	–0.44
Lead 1	Defocused attention	0.25	**0.67**	–0.42	–0.35	0.18	0.05
Lead 7	Engagement	0.01	**0.64**	–0.32	–0.19	–0.12	–0.05
Lead 4	Evocative activity	–0.08	0.56	–0.26	–0.10	–0.12	–0.13
Lead 5	Decoding	–0.29	–0.39	0.23	0.35	–0.30	–0.03

Note. The data come from correlations based on the contribution of each subjective variable (rating, emotion selection) and each EEG variable to the 20 different elements of the yogurt package. The "presumed meaning" of the leads in parentheses is used only for illustrative purposes, and comes from other work by the author.

RDE, with its systematic approach tying the different measures to specific elements that are manipulated experimentally, reveals the co-variation of the EEG response of a specific lead with the subjective measure of interest and emotion/feeling. This analysis can only be done by ***partialling out** the variability due to respondent and measurement time*, and then distilling the data to the contribution of each of the 20 elements. Through the intermediary of the 20 elements, each of which generates a pattern of EEG responses *and* a pattern of subjective ratings, we can correlate EEG with interest and with emotion/feeling.

SOME OBSERVATIONS

The goal of this experiment was to understand some of the aspects of the EEG as it relates to business-relevant stimuli. As such, we should look at highlights of the study, not so much for hypothesis testing, but rather as directions for neuromarketing research.

1. Experimental design helps. The analysis was made much more productive by looking at how one variable (feature of package) drove both subjective and EEG responses. Afterward, the relation between the subjective and the EEG response emerged more clearly. Lesson learned—look for lawfulness between stimulus and response

(controlled package feature *vs.* rating and controlled package feature *vs.* EEG; not between response and response or rating and EEG emotion).

2. Reduce inter-respondent variation. The variability due to respondent and measurement time had to be partialled out in order to make sense of the data. There is a lesson here. In order to partial time out, high resolution of measurements in time is needed.

3. Too much averaging can destroy the signal. By averaging the EEG readings in time to generate a single EEG value, the data tend to lose any meaning useful to the marketer. The longer the averaging interval, the greater the loss. Therefore, if we want to calculate an overall effectiveness coefficient for a marketing message, such as a commercial, a package, or a printed ad, averaging the EEG indicators during the whole duration of the message is the wrong way to go.

4. Some EEG leads are better than others. If the connection between the lead and the emotion can be validated, then furthermore, negative emotions are worth looking into. They may be a "treat" for the neuromarketer. Specifically, lead #8 (presumably linked with the emotion of "anxiety") appears to be promising. Lead #8 co-varies with three emotions at the same time: curiosity, boredom, and irritation. Want to find out how curious the consumer is? Measure the consumer's response using lead #8.

5. There are promising, but not conclusive, data about purchase intent. Further research is needed. Only lead #2 (focus) gets close to estimating purchase intent, with a correlation of –0.56. The lower the reading from lead #2, the higher the purchase intent.

6. Overall, lead #3 (memory storage) appears to be the most desirable state of consciousness of the consumer if we want to boost sales. It is associated to curiosity (correlation 0.65), excitement (correlation 0.45), and most important of all—interest (correlation 0.50).

7. The delta frequency band (lead #4, assumed to co-vary with the evocative factor) does not predict any emotions or purchase intent.

CRITICAL ISSUES, REVISITED

Let's review "critical issues" listed earlier in the light of the pilot project findings:

1. EEG is capable of showing with a certain degree of accuracy if the consumer viewing or listening to a marketing message is experiencing defocused attention followed by focus, anxiety, memory storage, and so on.

   Should we care? Yes, we should. We want the consumer to experience positive emotions, such as curiosity and excitement, and avoid negative emotions, such as boredom or irritation, while being exposed to a marketing message. In contrast, Table **3** shows that these emotions are significantly correlated to some of the EEG indicators already popular among neuromarketing researchers.

2. Can EEG indicators predict purchase intent?

   Not yet, but the results give some promising signs. The first four indicators in Table **3** have some sort of correlation with purchase intent. They don't have as strong a correlation as we would like, which would be 0.60 or above, but the results are promising.

For the time being, it appears that an innate mechanism prevents intruders from finding out what our purchase intentions are. We have to carry out a deliberate action in order to disclose our purchase intentions: our ratings. The free will that guides our decisions appears to be extended to the disclosure of our decisions as well.

BEYOND THE PILOT PROJECT—COMMERCIAL APPLICATIONS

Up to this point, we have used RDE to answer questions about neuromarketing, with the (correct) assumption that RDE has a long and solid record of successful predictions. What's next? Why should we use a new market research methodology (neuromarketing) if RDE already solves the same problems that neuromarketing is supposed to solve? This is the question that a skeptical reader might ask.

The combination of EEG with eye tracking, with both plugged into the RDE platform—as we will later see in this section—opens the path for the creation of a high-definition version of RDE.

EMOTIONAL NUANCES

For the pilot study, we selected five emotions/feelings that we felt to be appropriate to the yogurt study. It's worth a short digression to list the reasons why we did five, not more:

1. The more emotions we consider in an RDE study the harder the task will be for the respondent.

2. A respondent would be instructed to select ONE of the emotions/feelings from the set. The more options that we present to the respondent the more times an option would be left unchecked. It was important to populate the matrix of selections. Five options would allow a more populated data matrix for analysis by regression modeling.

Having found a strong correlation between most emotions (four out of five) and EEG indicators, opens the way to an ongoing, iterative research. The next time we conduct another research study, similar to the pilot study, we can exclude these four emotions from the survey: curious, bored, irritated, excited. The EEG measures are good enough for estimating those four emotions, and we can replace them with four other emotions. After a number of research studies, we will have many more emotions spotted on the EEG map.

Once we obtain a satisfactory richness of emotional nuances, there is virtually no limit to the level of detail we, as researchers, can have insight into. Using the yogurt package example, we would be able to associate each element to a long array of emotional nuances. We would not only be able to say that chocolate generates low interest—something that RDE easily found during the pilot project—but what can be attempted in order to increase its appeal: different image color, different shape, different size, different phrases about chocolate, *etc.*

Using a similar approach in copy testing, we can fine-tune the emotional impact of advertising.

For example, EEG synchronized with eye tracking helps identify those characters of a TV commercial—people or objects—that are deleterious or beneficial. Indeed, by studying the correlation between areas of interest looked at, and EEG indicators, it is possible to identify a pattern of emotions emerging any time specific characters or objects appear on the screen. Similar attempts have already been made with the aid of graphic tools. However, the level of accuracy and scientific confidence reached by RDE's statistical analysis is far better than intuitive, graphic observations.

Similar applications of RDE combined with EEG and eye tracking can be used for boosting the sales generated by product placement (advertisements of products embedded in television shows, movies, video games, and books) and by in-store marketing messages.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

The pilot project has been designed and executed by a team of US and European research groups:

- Rule Development Experimentation: Howard R. Moskowitz, Alex Gofman, and David Moskowitz, Moskowitz Jacobs, Inc., White Plains, NY, USA

- EEG: Francesco Gallucci and Riccardo Trecciola, 1to1 lab, Milan, Italy; EEG devices from Thought Technology, Montreal, Quebec, Canada

- Eye tracking: Johanna Fyrbjörk, Ipsos Sweden, Stockholm, Sweden; Niall Bellabarba and Mihkel Jaatma, RealEyes, Oxford, UK; eye-tracking devices from Tobii Technology AB, Danderyd, Sweden

- Project coordination: Sokol Zace, Contact Design, Boston, MA, USA

- Business applications adviser: Steve Onufrey, iNovum Philadelphia, PA, USA

REFERENCES

Bloom, F. E., Beal, M. F., & Kupfer, D. J. (2002). *The Dana guide to brain health*. New York: The Free Press (a division of Simon & Schuster).

Brat, I. (2010, February 17). The emotional quotient of soup shopping. *The Wall Street Journal*. Retrieved May 12, 2010, from http://online.wsj.com/article/SB10001424052748704804204575069562743700340.html?mod=loomia&loomia_si=t0:a16:g2:r4:c0.0537001:b30779526

Cade, C. M. & Coxhead, N. (1989). *The awakened mind biofeedback and the development of higher states of awareness*. Shaftesbury, Dorset, UK: Element Books.

Moskowitz, H. R. & Gofman, A. (2007). *Selling Blue Elephants: How to make great products that people want before they even know they want them*. Upper Saddle River, NJ: Wharton School Publishing.

NeuroFocus. (2010). *Direct measurement: Attention, emotional engagement, memory retention.* Retrieved May 12, 2010, from http://www.neurofocus.com/uk/neuromarketing.htm

Razumnikova, O. M. (2007). Creativity related cortex activity in the remote associates task. [Electronic version]. *Brain Research Bulletin*, *73*(1–3), 96–102.

Sands Research. (2009). *What can measuring brain waves tell us about an ad's effectiveness?* Retrieved May 12, 2010, from http://www.sandsresearch.com/AT_SRI_Quirk.aspx

Sederberg, P. B., Kahana, M. J., Howard, M.W., Donner, E.J., & Madsen, J. R. (2003). Theta and gamma oscillations during encoding predict subsequent recall. [Electronic version]. *The Journal of Neuroscience*, *23*(34), 10809–10814.

Wise, A. (1995). *The high performance mind: Mastering brainwaves for insight, healing, and creativity*. New York: Tarcher/Putnam.

CHAPTER 23

Balancing People's Future Demand and Design Genius

Marco Bevolo*

Independent Author and former director, Philips Design, Einhoven, The Netherlands

Abstract: Design is a multidisciplinary field in rapid evolution. For the purpose of this chapter, we focus on the pragmatic reality of corporate design teams. How do corporate design teams cope with the challenge to bring research onto their radars? Firstly, by being more and more integrated in the overall branding, marketing, and research processes of their respective corporations. If industrial designers of the past were accustomed to be in a dialog with R&D scientists in order to be recipients of technological innovation, nowadays corporate designers engage in a vibrant ongoing discussion with more stakeholders, both internal and external, representing the customer's voice within the company setting. In the past two decades or so, the role of research within corporate design processes has increased to the extent that entire departments were created almost by fiat. This was undertaken with the specific purpose to organically grow new competencies from within new portfolios of nontraditional and complementary design domains. Such approaches as ethnography, laddering, and future studies are among the specific approaches that have been increasingly integrated into the corporate design portfolio. The incorporation of these disciplines led to some exceptional cases of bottom-up excellence, as well as to some oddities in corporate portfolios and management directives. The new economic climate demands a rational approach to systematically anticipating people's needs and wants. Given today's drive toward fact-based decision making, the introduction of statistics and scientific methods appears to be the natural next step to streamline design skills. These methods should enhance innovation and delivery, and further encourage research competencies within the design portfolio of corporations.

Keywords: Creative leadership, design, design research, future study.

THE NEW ROLE OF DESIGN IN A POST-POSTMODERN SOCIETY

Perhaps no other corporate domain has enjoyed the same degree of visibility as design did in the decade before the fall of the Lehmann Brothers in 2008. Since the launch of products like the iMac by Apple, back in 1998, designers have found

***Address correspondence to Marco Bevolo:** Independent Author and former director, Philips Design, Einhoven, The Netherlands; Tel: +31 40 243 2034; Email: marco.bevolo@gmail.com

Alex Gofman and Howard R. Moskowitz (Eds)

the way to reach the hearts and seduce the minds of consumers worldwide. The first signs have been the relatively plain, yet revolutionary aesthetic developments like the introduction of innovative color schemes for just about any thinkable consumer product. In a wider context of analysis, designers managed to rise to the role of corporate innovators, thought leaders, and sometimes even "media gurus". At its climax, the popular love for design translated into the charismatic presence of designers, from Stefano Marzano, CEO of Philips Design; to Yves Behar, creator of the $100 PC for developing countries. The media has celebrated design as well, in key articles in magazines like *BusinessWeek*, in their innovation and design-focused regular issues, and even on the front cover of business publications like *Fast Company* in their "Masters of Design" issue of October 2007.

The "rational" approach to design of traditional European descent, from Bauhaus to the visions of former futurist, Bruno Munari and other design thought leaders, shifted to the postmodern and post-postmodern visions of new thinkers like Platform 21 and their "Repair" manifesto (Bevolo, 2010). This process of intellectual optimization was matched by the diverging trends of the collecting fever of DesignArt (Bevolo, 2010) and of sustainable focus in search of new consumption models. From designers aiming to serve society and progress, to designers acting as fine artists and media "prima donnas," and on to the ultimate status symbol of cultural achievement: "the" designer was seen as a new oracle of wisdom in our de-ideologized societies. Design ruled. The time for design is now, and design still rules in terms of popular culture and collective beliefs. Ideologically speaking, design is one of the value paradigms in the new "weak ideologies" of our "liquid societies" (Branzi, 2006).

What is left of the connection between customers and plain products? The rise in popularity of events like design fairs and shows may lead us to lose our focus that the business of companies that produce everyday objects and packaged goods is done so within a standardized production chain, and for financial goals. Can we discover a connection between the work of the new creative gods of design and what "the rest of us" truly want?

The answer to the foregoing is "yes", but a "yes" that is not very clear to most people. Below the surface of apparent glamour and sometimes excessive selfcelebration, designers in general, and corporate designers in particular, actually have been busy with the constant redefinition of their profession. In the past two decades, different corporate design departments and more “heavyweight” design firms operated with different methods, but with one goal in common—a general drive to study the human mind, the human body, and human emotions by means of systematic research.

DESIGN AND RESEARCH: GENESIS OF A GREAT LOVE, WITH SOME MINOR MISUNDERSTANDINGS

The name of the game for the best of corporate design in the past two decades has been to *understand people*, with research rigor to support creative excellence. One might honestly state that across the 2000s every “serious” design entity positioned itself as being “research led”. What changed is the actual nature of such research, from ethnography and immersion to trends and experimentation. Here, different approaches could be followed and tracked over the years.

One good example of this approach is IDEO. A leader in the setting of the design agenda, IDEO is an independent firm, although for several years it was part of the Steel Case “galaxy” of assets. Indeed, IDEO is not counted among corporate design players. However, its methods and visions have strongly influenced the everyday practices and ambitions of corporate design centers worldwide. An accurate media relations strategy and a skilled thought leading stream of publications (Kelley, 2001) make IDEO by far the key US design firm worldwide, especially when it comes to focusing on innovation and thought leadership (Kelley, 2005). In retrospect, one might state that IDEO opted for an “action-oriented” research approach. For example, this meant sending designers as “research agents” in the very field of action, and letting them experience for a couple of weeks the everyday life of those very people for whom they would be designing. It might be possible to recognize in this way of working a potential reference to microsociological approaches from the likes of Erving Goffman. In reality, as it often happens with designers, the plain truth might be much less academic, and much more pragmatic. IDEO took an approach to research geared not so much in science, but in practice: “doing” over “studying”. It surely worked in a number of great projects, and it greatly helped to steer the firm through the difficult biennium of the global recession with its leadership status intact.

IDEO injected the power of design into research, and *vice versa*. We are, however, still in the domain of traditional research techniques. The idea to perform "research through design" was the next step. This is yet another perfect example of how the whole creative industry field expanded its ambitions from plain aesthetic leadership to gaining a deeper understanding of people's lives. Here, leading educational institutions in the design sector, from the Royal College of Art of London to the Domus Academy of Milan, played a big role in seeding the idea that designers might "know" more about reality thanks to their intuition and talent—if the latter is systematically adopted as a research tool using appropriate methods.

An outcome of this "deeper ideology" of what design is and how far it can reach can be identified in several groundbreaking programs on future lifestyle scenarios (Marzano, 1995). Here, Philips and their 1995 "Vision of the Future" (Marzano, 1998) offered perhaps one of the first attempts to summarize such ambition, and bring it to actuality by means of a visual landscape of prospective technological artifacts. That is, Philips made the future design real. This was a vision based in the groundbreaking research work by Future Concept Lab of Milan (Morace, 1995). Morace and his team contributed with their matrix tools to sociologically envision the future that Philips designers materialized in forms and applications for people. Nevertheless, it might be concluded that the ultimate lead was not in the research rigor but in the creative talent of designers and in their "soft skills" to conceptually and visually render tomorrow's digital applications in appealing maquettes, leading to a trendsetting eye-candy formal language.

Complementary to this design approach, the idea emerged that the future is co-created with commissioners and with end users. The idea became increasingly popular among various creative industry and architectural firms: from AMO, the think tank by architectural guru Rem Koolhaas; to ARUP, the engineering leaders worldwide; back to IDEO. An increasing number of design-oriented studios saw the birth of new services in future innovation research. Here, in the early 2000s, IDEO marked yet another milestone in the industry by defining the format of card sets for future studies. This format was conceived in parallel by Philips Design as well (Vissers, 2005) and then at later stage adopted, with slight variations, by the likes of ARUP foresight practice. Within this approach, highly multidisciplinary workshop teams are confronted with simple, actionable game cards, reporting facts and trends, for immediate actionability while they are also defining future scenarios,. When correctly deployed, the concise sentences printed on such "gaming tools" would translate into triggers to provoke and shake participants, sometimes resulting in the facilitated creation of future scenarios of great power. At some point in the past decade, it became apparent to more and more stakeholders in business innovation circles that designers could manage not only to create the present, but also to anticipate tomorrow's evolutions (Bevolo and Price, 2006).

In time, the adoption of research capabilities ended up integrating design with other consumer-oriented components of the corporation's "knowledge" portfolio. These components ranged from market intelligence to brand strategy, and on to advertising, respectively. This strategic extension of design grew by including specialists in human factors, in product–service interaction, and even in colors and finishing.

Technical departments working on the next generation of textiles and coatings for product design specifications have been in existence within leading industrial creative firms like Italdesign Giugiaro since the 1980s. Here, technological and aesthetic roadmaps for pigments and other materials are constantly analyzed and integrated in the advanced phases of the industrial process. The conversion of these departments into bridgeheads of future research was one of the possibilities for corporate design centers aiming at the creation of a more strategic portfolio.

From these early seeds, the design environments of corporations such as Philips became the more or less natural hosts of implicit-research capabilities bordering on (and sometimes even overlapping with) other research departments. Unlike today, the economic context was there to help with this new enterprise. The natural optimism before and even after the Internet bubble resulted in a natural drive and even a thirst for visionary directions in terms of what the future might bring, especially opportunities derived from new digital applications (Aarts and Marzano, 2003).

During the late 1990s bubble, the whole field of future studies became very popular in the design industry: a number of corporations engaged in advanced design projects, with the likes of Motorola and Whirlpool creating their own visions of the future. In the larger context of the consulting market, the field of future studies permitted the flowering of a number of independent firms offering portfolios of trends in the form of aesthetic roadmaps, future insights, and other platforms in anticipation of the next lifestyle. Among these players, one might recognize the emergence of dynamic, small consulting firms like StreativeBranding of the Netherlands, Sputnik of New York, USA, or The Future Laboratory of London, UK. Historically, the common point of strength of these enterprises has been their ability to anticipate what consumers *will* want, and insert such information into the corporate processes of their customers.

The method adopted in this field of operations remained as vast as it was flexible. From cultural studies to more classic market research techniques, the heart of these players was with the consumers of today and tomorrow, but not with the precision of statistics. A natural host for the reports and consulting input from these new players was the design department of corporations. Here, the combination of visual appeal and visionary drive could strike a chord, which it did, often quite successfully.

DESIGN-FRIENDLY METHODS

Under the conditions that existed before the economic downturn, a purely qualitatively driven, designer-friendly approach seemed inspiring yet methods based. There was a great promise of results:

1. Extension of corporate design portfolios into new territories of innovation and research, with the benefit of a wider reach for their teams in terms of consumer understanding.

2. Efficient, effective "plug in" of contributors to the strategic processes within the corporation. These contributors could be internal (*e.g.,* the design department) or external (*e.g.,* consulting firms). These contributors allowed corporate players to explore new methods, some of which were scientific, others of which were of a somewhat "pop" nature when viewed from the scientific viewpoint.

This overall blossoming of both internal capabilities and external consulting practices resulted in a highly stimulating environment for creative leaders of all sorts. These leaders generally relished being fed with "soft data," mostly of a visual or narrative nature. Corporate design directors and their eager, motivated teams were, in turn, afforded new opportunities to form their own visions and then execute their ideas based on "informed intuition". This blossoming took place within a welcoming yet not-too-critical context, where scientific facts were sometimes disregarded as nonrelevant. It seemed, at the time, as if a perfect balance between art and science, theory and practice, hope and reality, had been achieved. History would show, however, that such a heralded achievement was more a fata morgana effect than a true fact.

CO-CREATION *VS.* DESIGNART: TWO EXTREME OPPOSITES OF YESTERDAY'S FUTURE OF DESIGN

Let us quickly review the context where the natural drive of corporate design toward futures research took place. It is appropriate to take a step back and focus on high-tech companies and their own design capabilities, from the "800-lb. gorilla" (Microsoft) to the smallest start-ups in emerging countries. Parallel to the aforementioned trend analysis, a strong capability toward co-creative design was introduced by digital platforms and digital technology. The Internet and its design applications opened the creative process to users in unprecedented ways.

The protocols of the corporate design process changed its very nature: not just the practice of anticipating people's values as based on research reports or multidisciplinary workshops but also an actual dialog with people in real time. Co-creation meant the opening of the design profession to people, with the active and concrete intervention of these newly invited people in the creative process. Here, new research-based start-up firms such as Experientia of Turin, Italy, managed to redefine the way research connects to the creative process, showing the way to corporate players like Vodafone or Nokia. "Co-design" became a mantra for many. Yet its true implementation remained somewhat anecdotal, mostly lip service for most. It is undeniable that a number of established and emerging designers polarized the creative services market on the basis of their unique individual charisma. These "personal brands" of design leadership followed a line of creative production mostly based on their own vision, hence disregarding the viewpoints of co-creative outsiders in their everyday work. As much as this phenomenon might look like the opposite of co-design, such an approach to design direction actually responds to deep human psychology. "Charisma" is a driver rooted in our social mechanisms. It represents an important segment of our very own collective DNA, whether coming from populism or from true leadership.

The rise of the "DesignArt" in the fine arts market helped to amplify this worship of design, but at a price. Corporate designers were confronted with two conflicting role models. On the one hand, we find the co-design facilitator. On the other hand, we find the Renaissance "maestro" of DesignArt. It seems quite peculiar that while communities of designers were experimenting with digital solutions specifically created to put users in the lead, communities of collectors were bringing the notion of design back to pre-modernist standards. As much as the DesignArt phenomenon was studied as an "evolution" of the (premium) markets, it appears now in retrospect to be a reactionary regression of design toward politically conservative, elitist "styling". To some, DesignArt represents the actual nemesis of what democratic forms of "open source" creation aim to be.

In recent years, corporate design departments experienced growth and prosperity as never before. Co-design was the "democratic" side of their future, offered through digital applications and by enlightened marketers. Co-design represented a territory that was not yet perceived to endanger the status of the corporate designer. DesignArt was, at the same time, the glamorous spot of aspiration that secretly seduced every corporate designer in terms of their own self-perception. However, the crisis came. Under the new economic circumstances, the drive to base the entire corporate life on facts emerged as one of the new imperatives. This traumatic and sudden change of scenery set the clock back to its starting position, bringing in a new set of brutal priorities; reality kicked in, replacing dreams of democratization or glory. To summarize our exploration so far, by the time U.S. President Barack Obama was elected in 2008, the general landscape of interdependence between design and research prior to the downturn showed a clear consolidation:

- Most corporate design practices included some research capabilities within their portfolios. These capabilities ranged from human factors research to more advanced and integrated consumer and market research. We find a strong preference for qualitative techniques and socioanthropological themes.

- Most research capabilities were connected to specific design applicability, with aesthetic trends and ethnography as the most prominently disciplines. These forms of information were most easily adopted by multidisciplinary teams, given their relative ease of use and comfort of translation for designers.

- A rather rapid proliferation of outsourcing practices ended up establishing a flourishing supplier consulting base of small but dynamic agencies working in areas like specific industry trends or customer insights.

Just as an agency like IDEO extended the realm of “design” into microsociological analysis and similar exercises, multidisciplinary teams such as Philips Design Strategic Futures incorporated research in ways unknown in earlier decades. At the start of the new crisis, one could not realize the real need and value of scientific data with statistics. It was felt, as always, that intuition, not science, would lead to superior creativity. As the crisis deepened, however, it became increasingly clear that the old ways of pure intuition simply would not work as well, or in some cases not even work at all.

THE REALITY CHECK OF TODAY, THE POSSIBILITIES OF TOMORROW

Even before the latest economic breakdown, a reality check in this field of research exploration and application was long overdue. In spite of the efforts by enlightened management and committed individuals, one might conclude that within corporate design teams across all industries, “research,” for the most part remained a rather separate island chain of specific competences. As odd as it might sound, the above statement applies to most of the star research teams within corporate design departments in the 1990s and 2000s. What seemed a natural renaissance in terms of humanistic focus turned out to be an “implant” within the corporate flowchart of organizational systems. Even in the best of cases, research remains a clearly separate box in the corporate process of design, sometimes operating in antagonism with marketing, market intelligence, and other departments.

The dramatic emergence of the post-downturn world brought an unexpected side effect: the need for hard facts. Prompted by the crisis, this deep need for facts disrupted this fragile balance point between designers and their own internally allocated researchers, in two ways:

- Confronted with the need to downsize, corporations did not exempt their design departments. On the contrary, research teams were often identified as being non-mission-crucial competencies in design portfolios. Hence they were vulnerable to elimination from the corporate structure.

- The continuity and the effectiveness of qualitative research combined with the designer's intuition were only sometimes apparent. Yet the connection was not sufficiently strong to justify the related corporate investments. At the end of the day, what appeared actionable to the designer remained vague and ineffective to the corporate VP of marketing or corporate strategy.

As a result, one might say that the ultimate effect of the recent crisis on the corporate design sector was its general retrocession to the status quo before 1989, at least from the viewpoint of the integration of market research within a corporation's master processes. Of course, such status is not sustainable over the long term. Just as in consumer markets, the clock of corporate design cannot be set back to earlier stages of its evolution. The next challenge for all players and stakeholders involved in this process will then be how to move the "corporate design research" clock forword to 2012, and even more importantly, how to start truly looking into the future?

A CONVERGENCE OF PARALLEL LINES: RDE AS A CREATIVE ASSET FOR DESIGNERS

As much as design provided (sometimes glamorous) solutions to meet people's dreams and demands in the past 10 to 20 years, the world has profoundly changed, and in some cases has "moved on". As Steven van den Kruit, creative leader and successful researcher of change within aesthetics and cultures at Firmenich, Geneva, stated: “Consumers already made the U-turn”. The change in markets and societies has been deep, sudden, and truly definitive. In contrast, corporations and their research departments are sometimes still frozen in the no-man’s land between their past and their heavy cost-cutting and restructuring. It is now up to designers to find the necessary new ways, new methods, and new connections between research and creative leadership to meet the challenges of today, and to anticipate the possibilities of the post-downturn, whether recovery or even perhaps lack of recovery.

The next steps in the dialog between (corporate) design and (future-oriented) research will require a clear response to a very dense cloud of challenges:

1. The scientific rigor of a research solution must be demonstrated now, in a way that it did not have to before. The times of vague trend directions delivered by self-proclaimed “gurus,” *e.g.,* in the fashion of Li Edelkoort in the 1990s, appear to be over. Companies need to base their own decisions on facts, not opinions.

2. The need for immediate applicability of research solutions will militate against complicated, complex, incomprehensible statistics that impress but do not enlighten. Designers will not take on board complicated and complex statistics. Their hearts are, and remain, centered around the purity of creative direction, where they expect to exercise their own measure of charisma.

3. Including end users within the design process will continue, and increase. As discussed above, consumers’ expectations will not be less important, but rather more important. The consumer as co-designer is becoming increasingly acceptable in advanced economies and is rapidly spreading to the thought leaders of BRIC (Brazil, Russia, India, China) countries as well.

SYSTEMATICS USING RULE DEVELOPING EXPERIMENTATION (RDE)

Rule developing experimentation may constitute one of the ways by which the designer can blend intuition, talent, consumers, and rigor. RDE could be viewed as global approach to systematic design that can connect together the different aspects of cultural developments in the leading cities of the world and deliver the highlights to the benefit of designers, who can act upon the specifics of those developments. RDE combines the granularity of development with the people-centered nature of information (always critical for designers), and merges in the expected "dollar value" of the different aspects, to guide the combination of design and commerce (Bevolo *et al.* 2009).

The principles of RDE, articulated in the case history of "Always On" at Hewlett Packard (Moskowitz and Gofman, 2007), show that RDE applications fit the design initiative quite well. The RDE information, expressed in the impact of different specific ideas, combines the precision of quantitative data with the flexibility of "plug and play" information streams. The result is a useful stream of information simultaneously flowing from the consumer and society across to R&D, to engineering, and to design departments, respectively. With RDE, end users open the design process with their direct and immediate input, while at the same time company departments are unified by the power of shared insights and information in terms of the same data sets.

There are always problems, of course, and when we deal in the world of design, many of these problems are "cultural". To state that quantitative solutions will be warmly welcomed in the heart of design centers would be a bit self-deluding, overly optimistic, and yet is to be desired Over time, and in today's business environment, the pressure from the world outside to "open" the design process by means of co-creative practices will only increase. RDE and other worldviews of this type will turn into a parallel road map for knowledge. These solid, science-based yet art-sensitive approaches will address today's increasingly dramatic demand from top management for hard facts useful in decision making. All signs therefore point to an increase of demand for designer-friendly quantitative solutions.

RDE offers one additional advantage of particular importance to the world of artistic, yet functional, design. Although a statistics-based method worldview with a full-blown "scientific pedigree," RDE was born from genuinely curious minds coping with the granularity of everyday life. RDE thus becomes a tool for the designer, merging as it does one's life with scientific rigor.

What we need in this brave new world of tomorrow is the ability to think and to evolve by crossing borders. It is inevitable that we will move increasingly faster toward a global business environment where individual disciplines will simply not be enough. Hence, even the best mathematician or the most talented designer will fail, if the person operates in splendid isolation. It will be up to the next generation of designers to rethink design practices by capturing the essence of charisma, and yet (here's the hard part), also leaving behind any ego dynamics and artistic pretensions. It will be up to tomorrow's designers to manage a process that returns to its roots: a modernist and optimist view of what design can do. It will be up to the future generations of RDE developers to take on board the rational and emotional demands of designers, and increasingly include these demands in the functionalities and features of their knowledge-building methods, which will be to the ultimate delight of creative communities.

To those who superficially skim the two surfaces of rule developing experimentation and corporate design, the vision of joining together knowledge and design might not look like an easy dialog to initiate, much less a marriage with a promise of enduring. In reality, the most advanced explorers of these two separate continents already established a number of bridges and bridgeheads at theoretical level and with practical projects. Such bridges could be erected in a variety of areas:

1. To understand end users and future consumers of products and services. Through the marriage of RDE and design, companies will design better, faster, and more inclusively of people's dreams and wants.

2. To satisfy the market demands. The direction has already been set here. The increasing acceptance of statistical-based experiments within the structural flow of corporate design processes will be accelerated by the pressure of markets and customers, and by the sensibility of all involved stakeholders. Competition is the friend of knowledge, the bane of prima donna aloneness.

3. Economic "corrections" drive advances, and shake out old shibboleths. In the adversity of the economic downturn, this one

process of dialog and integration between science and creative leadership is one of the most promising motivations of a new optimism, one grounded on both hard facts and visionary talent.

CONFLICT OF INTEREST

None declared.

ACKNOWLEDGEMENTS

None declared.

REFERENCES

Munari, B. (1971). *Artista e Designer*. Roma/Bari: Laterza.
Bevolo, M. (2010). *The golden crossroads*. London: Palgrave Macmillan.
Branzi, A. (2006). *Weak and diffused modernity*. Milan: Skira.
Kelley, T. (2001). *The art of innovation*. New York: Random House, Inc.
Kelley, T. (2005). *The ten faces of innovation*. New York: Random House, Inc.
Marzano, S. (Ed.). (1995). *Past tense, future sense*. Amsterdam: Bis.
Marzano, S. (1998). *Facts + thoughts*. Amsterdam: Bis.
Morace, F. (1995). *Metatendenze*. Milan: Sperling & Kupfer.
Aarts, E. & Marzano, S. (Eds.). (2003). *The new everyday*. Rotterdam: 010.
Moskowitz, H. & Gofman, A. (2007). *Selling Blue Elephants*. Philadelphia, PA: WSP.

CONFERENCE PAPERS

Vissers, B. (2005). *From aesthetic trends to new value signs*. ESOMAR "Digital Futures" Conference, Paris.
Bevolo, M. & Price, N. (2006). *Towards the evolutions and revolutions in futures research*. ESOMAR "BrandMatters" forum, New York.
Bevolo, M., Shofu, R., Moskowitz, D., & Moskowitz, H. (2009). *Putting fragrance into context*, In Proceedings of ESOMAR Fragrance Conference, Cannes.

END NOTE

One of the better-known Albert Einstein aphorisms is "*Imagination is more important than knowledge*". The ancient dichotomy between what we know and what we imagine or sense by instincts is relevant in almost every domain. Yet, what if this truism could be revised, so that imagination and knowledge were combined? The outcome might be far much more powerful, exponentially and synergistically enhancing the contributions of its individual elements. The possibilities may be dramatic, and quite possibly help us reach far beyond where we are today.

Mind Genomics® (MG) and its knowledge development tool, Rule Developing Experimentation (RDE), address the challenges of how people react to their changing and increasingly interdependent daily world. Mind Genomics® is founded on both imagination and knowledge. Its overall goal is to dissect specific, everyday experience into components, get people's responses to these components, the dimensions of the everyday, identify different mind-sets of people defined by how they respond to the aspects of the everyday, and then organize the information into accessible databases. All of this is the grand plan, to understand how the mind responds to its world of experience, treating the information as genomes of the mind.

Like a chef who uses common food ingredients like most of us, but is able to create an exquisite meal, the combination of MG and RDE unveils new and plentiful information, made out of the stuff of the everyday. Like a talented chef, MG and RDE open up new domains, taking us closer to better consumer comprehension through a basic understanding of experience. The "delightful" output is solid data, actionable information, fact-based insights about what drives decisions, and the ability to better address the needs and expectations of consumers and/or customers.

Traditional mind-sets and galloping resistance are the major hindrances to new ideas and their utilization. No doubt, Mind Genomics® will meet such resistance, for it breaks new ground; but not to fear. This book offers a wide scope of topics that were studied using MG and RDE. The book brims with in-depth knowledge

Alex Gofman and Howard R. Moskowitz (Eds)

and personal experiences. The novice and the more advanced reader should benefit from the depth of information. Hopefully, much of the resistance to the methods will melt away when readers confront the solidly empirical aspects of the book and the wealth of data.

The book covers a wide spectrum of topics including: theoretical foundation of MG and RDE, optimization of practical sensory, message and structured package and website, advertising research, brand communications and experimentation and consumer-driven innovation. Like an idea that represents a series of new connections between neurons in the brain, the book should serve first and foremost as a stimulus for more knowledge, more intellectual adventure. The sheer variety of topics offers the reader that delightful serendipitous environment which promotes collisions of ideas, and excites innovation.

What made this book unique? Simply, this is a seminal book. It unveils a new approach to identify general patterns of what excites consumers. We predict that in the near future the MG and RDE ability to delve into the human mind will be combined with brain studies (*e.g.,* functional magnetic resonance imaging, fMRI) for better insights about what drives decisions. The possibility to study consumer behavior and identify the location and pattern of major brain activity in the cortex and other structures is, and will continue to be, extremely powerful. Perhaps most exciting news is that at least some regions of the brain continue to generate new neurons in adulthood, and those neurons appear to participate in the learning and memory process. MG and RDE should play a role in making this new research more disciplined, and better linked to the experience of the everyday.

In the opinion of this writer, this edited book, assembled by two well-known researchers, with contributions from experts in business who share a passion for applied science, has made a significant impact on the immediate domain of consumer science and behavioral research. But there's a beyond, a world of application of Mind Genomics® and Rule Developing Experimentation to the bigger world, the world of education and social policy. Those applications remain for the next volumes. The efforts of Gofman and Moskowitz and their contributors have, with this book, given the world new avenues, new opportunities to attract experts in various domains, whose contributions will pull the field forward. This book provides all of us with a newly sprouted garden where knowledge and imagination mix and flourish.

Sam Saguy, D.SC.
Professor of Food Technology and Innovation
The Hebrew University of Jerusalem
Israel