

# Big Mind + Big Data: Empathy and Experiment build Mind Genomics

Howard Moskowitz, Nadya Livshits,  
Linda Ettinger Lieberman, Dvora Chaiet,  
David Moskowitz (*Authors*)  
Research Department  
Moskowitz Jacobs, Inc.  
White Plains, NY, USA  
mjihrm@sprynet.com

Batool Batalvi (*Author*)  
S B & B Marketing Research  
Toronto, ON, Canada  
Lahore, Pakistan  
batoolbatalvi@gmail.com

**Abstract**— This paper presents a new approach to Big Data. It introduces the notion of Mind Genomics® as a way to profoundly understand people, not so much by their mind as by the pattern of their reactions to messages. Understanding the way *nature is*, however, does not suffice. It is vital to bring that knowledge into action, to use the information about a person’s mind to drive behavior. The paper introduces the Viewpoint Identifier as that tool, and shows how the Viewpoint Identifier can move to scoring entire databases. The paper closes with the vision of a new web, Big Mind + Big Data, where the networks developed show both surface behavior that can be observed, and deep, profound information about the way each individual *thinks* with respect to a variety of topics.

**Keywords**— *Big Data; Mind Genomics; reactions; behavior; Viewpoint Identifier*<sup>s</sup>

## I. INTRODUCTION

When we look at networks, seeking patterns, we infer from the behaviors and the underlying structure what might be going on in the various nodes. We don’t actually communicate with the nodes; they’re represented geometrically as points of connection. Analytically, we can look at behavior, imparting a structural analysis to the network, looking at the different connections—the nodes, the nature of what’s being transacted, and the number and type of connections. By doing so, we infer the mind of the node. But, what about that mind? What do we know? What can we know? And more deeply, is that mind invariant, unchangeable, reacting the same way no matter what new configurations of externalities are imposed?

These are indeed tough questions. The scientific method teaches us to recognize patterns, regularities, and from those patterns to infer what might be going on at the origin, at the object doing the action, at the object lying at the center of the connections. Mathematics furthermore unveils these networks, different patterns, in wonderful new ways, highlighting deeper structures, and revealing hitherto unexpected relations. Those lucky enough to have programs with false colors see the patterns revealed in marvelous reds, yellows, blues, and greens, like the picture of stars of the universe. The underlying patterns are clearly not in color, and the universe does not appear to us so

comely and well colored. It is technology which colors and delights us, technology which reveals the secrets.

Now for the deeper question. What is beyond the network, the edges, inside the nodes, inside the mind of the center of the connections? Can we ever interrogate a node? Can we ever ask a point on a network to tell us about itself? Does the point remain the same when we shift topics, so the representation is no longer how the nodes interact on one day, but rather interact on another day, or in another situation?

## II. HOW TO THINK ABOUT THE PROBLEM

### A. Moving from analysis of an object to interrogating it

We move now from analysis of an object in a network to actually *interrogating* the object in order to understand it from the inside, to get a sense of its internal composition. The notion here is that once we understand the network as externalities *and* understand deep *mind* properties of the nodes in the network, the people, we have qualitatively increased the value of the network by an order of magnitude. We not only know how the points in the network, the people, react, but we know correlates of that reaction, the minds and motivations of these points that are reacting.

Just how do we do that when we recognize that this mind may have opinions, that the mind may have a desire to be perceived as politically correct, and that, in fact, this mind in the object may not be able to tell us really what’s important? How do we work with this mind to find out what’s going on inside?

It is at this juncture that we introduce the notion of Mind Genomics, a metaphor for an approach to systematically explore and then quantitatively understand the inside of the mind of a person or node in a system. The output of that understanding comprises content (the components of this mind), numbers (a way to measure the components of the mind), and linkages (the assignment of the content and its numbers to specific points, nodes, people in the network) [1,2].

B. A worked example – what should the financial analyst say to entrance a prospect to commit

Lest the foregoing seem to be too abstract, too esoteric, too impractical, let's put a tangible aspect onto the idea. What happens when the point or node corresponds to a person walking in to buy a financial retirement product from a broker whom the person has never met? How does this new broker understand what to say to the person at the initial sales interaction, that first moment of truth when there is a chance for a meaningful purchase to occur? And what happens when the interaction occurs in an environment where the financial consultant or sales person never even meets the prospective buyer, but rather relies upon a Web site, or a simple outward bound calling center manned by non-professionals? We have our network, nodes connected by the sales activity. By *understanding* the mind of the prospective customer, the financial analyst has a much greater chance of making the sale, in contrast to simply by knowing the age, gender, family situation, income, and previous Web searching behavior—of the prospect, all available from Big Data and grist for the analytic mill. We want to go deeper, into the mind of that prospect.

III. UNDERSTANDING THE “MIND IN THE NODE”

A. What drives a person to do something

Psychologists and marketers have long been interested in understanding what drives a person to do something, the former (along with some philosophers) to create a theory of the mind, the latter to create products and services, and sell them. We know that people can articulate what they want, describe to an interviewer the characteristics of a product or service that they would like, often just sketchily, but occasionally in agonizing detail. And all too often this description leads the manufacturer or the service supplier on a wild-goose-chase, running after features that people really don't want, or features which are so expensive as to make the exercise simply one of wish description rather than preparation for design.

A more practical way runs an experiment presenting the person, this *node* in the system, with different ideas, different descriptions about a product, obtains ratings of the description, and then determines through statistical modeling the particular elements in the description which link with a positive response. In other words, run an experiment treating this *node*, this point in a network, as a sentient being, not just as something whose behavior or connections are to be observed as objective, measurable quantities. Looking at the network as an array of connected minds, not connected points, minds with feelings, desires, and opinions, will enrich us dramatically in theory and in practice.

The experiment, or better the paradigm of Mind Genomics, is rather simple. We use a paradigm known as

Empathy and Experiment, empathy to identify the ‘what,’ the content, and experiment to identify the values, the ‘important’[3].

Our strategy is simple. We want to add a new dimension to the network by revealing the *mind* of each nodal point. To do so requires empathy, understanding the ‘what,’ and experiment, quantifying the amount, revealing the structure. Putting the foregoing into operational terms, we will identify a topic area relevant to the node, the person, uncover elements or ideas appropriate to the topic, and then quantify importance of each element. After Empathy uncovers the raw materials, the elements, Experiment mixes and matches these elements into different combinations, obtains ratings of the combinations, and then estimates how the individual elements in the combination drive the response.

The foregoing paragraph described an experiment., not a questionnaire. Rather, we infer what the person, the node, wants by the pattern of responses and from behavior we determine what elements produce positive responses and what elements produce negative responses [4].

B. Putting Mind Genomics into action

The best way to understand the concepts of Mind Genomics, its application to knowledge and to networks, is through an illustration. This paper presents the application of Mind Genomics to both understanding the micro-science of choosing a financial advisor for one's retirement planning. The case history shows the input and practical output of Mind Genomics, how a financial advisor can understand the mind and needs of his customer, identifying the nature of his customer from the very beginning.

Mind Genomics follows the general schematic, shown in Fig.1. Our focus is on selling financial retirement services through a broker who deals with customer prospects. Our customers are the nodes in the web who are to be approached in the way that is appropriate for each person particularly.

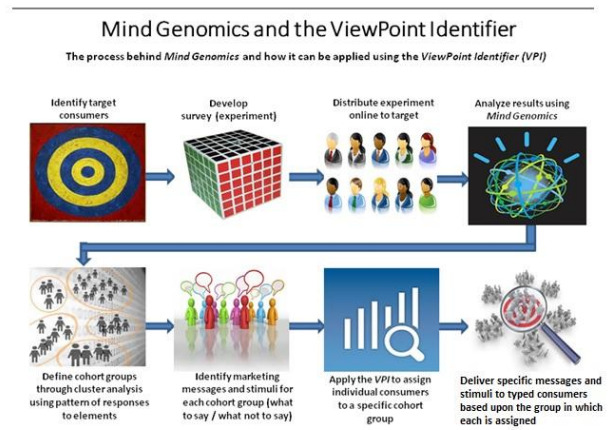


Fig.1: The process of Mind Genomics, from setup to analysis and application. Fig.courtesy of Barry Sideroff, Direct Ventures Ltd.

To create and to apply the micro-science we follow the steps below. Although the case history is particularized to selecting a financial advisor, the steps themselves would be followed for most applications. Only the topic area varies.

1. We begin by defining the topic area. We also specify the qualifications for the consumer respondents, those who will be part of what might initially look like a Web-based survey, but in reality constitutes a systematic experiment. For our study, the very specific focus is on the interaction of the financial advisor with the consumer, with the specific topic being the sales of retirement instruments such as annuities. The key word here is *focus*. Specificity makes all the difference, elevating the study from general knowledge to particulars that can be immediately applied in practice.
2. We develop the element (survey) matrix, the set of elements that we will use. Since our focus here is on the inside of the mind, what motivates the person to listen to the introductory *sales message* of the financial planner, we will use simple phrases, statements that a prospective client of the financial analyst is likely to hear from the analyst himself or read in an advertisement. Table 1 presents the set of 36 elements divided into four silos with each comprising nine elements. The elements constitute simple stand-alone messages painting a *word picture*. The silos act as convenient bookkeeping devices allowing us to ensure that elements that could possibly contradict each other if they appear together never do appear together. Thus, brands such as Wells Fargo, Prudential, Chase Manhattan, and Merrill Lynch, are all designated to the same silo, Silo A. When it comes time to create the test stimuli, elements from the same silo can never appear together. A set of 36 elements covers a great deal of ground and typically suffices to teach us a lot about the particular minds of the participants, our respondents or *nodes* in a web. The particular arrangement of four silos and nine elements is only one popular arrangement of silos and their associated elements. An equally popular arrangement is 6x6, six silos with six elements in each.

Table 1: The raw material of Mind Genomics, elements arranged into four silos, each silo comprising nine elements. The elements are the building blocks of ideas. The silos are merely bookkeeping devices which ensure that mutually contradictory elements do not appear in the same test vignette (concept).

Silo A	
A1	Wells Fargo...It is never too late to plan for your retirement
A2	Prudential...It is never too late to plan for your retirement
A3	Chase Manhattan...It is never too late to plan for your retirement
A4	Merrill Lynch...It is never too late to plan for your retirement
A5	Tell us when you want to retire and we will develop an action plan to get you there
A6	Assess your financial health with our retirement planning worksheet...answer some basic questions about your financial present and your goals for the future...then we will recommend some strategies to get you there
A7	We will work with you every step of the way to develop and monitor your retirement investments to ensure that they will meet your objectives

A8	We offer you pointers, recommended reading and worksheets so you can develop a retirement plan that suits your needs
A9	Your plan will be reviewed annually not just by YOUR consultant but with our entire staff of retirement planning experts...working in collaboration to ensure that your needs are being met
Silo B	
B1	Convenient 24/7 online access to your account and our experts...ask questions...monitor your progress...all online!
B2	We have not forgotten that we are in the relationship business...our retirement planning experts are never more than a phone call away
B3	Our philosophy is based on interpersonal relationships so we will want to meet with you and get to know you...this enables us to develop a better retirement plan for YOU
B4	As you near retirement...you can take steps to improve how long and how well your hard-earned savings will work for you during your retirement years
B5	During your peak earning years we will help you to manage the changes in your life to the advantage of your longer term retirement plan
B6	It is never too early to start planning for your financial future...the steps you take today will significantly affect the quality of your life 10,20 even 40 years from now
B7	We offer a "no surprise" fee structure...one flat monthly fee to establish and monitor your personalized retirement plan
B8	Our fee structure is based on the premise that you only pay for the services you use
B9	We offer a combination fee structure...a low monthly fee plus discounted prices for each transaction
Silo C	
C1	With you in charge of your retirement plan...you know if your objectives are being met
C2	Planning for your retirement now means you won't get caught short since you still have time to do something about it
C3	If you don't plan for your retirement...nobody else will...make sure you get to where you want to be
C4	A retirement plan will ensure that you have done all that you can to retire when you want and live the lifestyle you want
C5	Your retirement nest egg may be your single largest asset...invest it wisely with our certified retirement planners
C6	Work with one of our experts to develop your own retirement plan...because Social Security was never meant to be a full retirement program
C7	You will feel more secure knowing that every step of your financial future is being managed
C8	Our Retirement Resource Center has the tools, information and experts to help you make smart retirement planning decisions
C9	By drafting a retirement plan you will be securing your own financial future
Silo D	
D1	You can trust that your local retirement planning expert will be around as you near retirement...they won't be part of some impersonal conglomerate
D2	You can't underestimate the value of your retirement planning expert knowing you, your family, your community
D3	With a nationally recognized retirement planning firm you are never far away from expert advice...no matter where you travel...business or pleasure
D4	A special class of service is available to those who attain Platinum status
D5	Becoming a member of our Platinum Club says to others that you have "arrived"
D6	For a limited time only, meet with one of our retirement planning experts for a free initial consultation
D7	Available exclusively to our VIP customers, a complimentary consultation with someone from the Wills and Trusts department...ask the questions you have been meaning to ask
D8	Planning for your retirement has never been more convenient
D9	Take our Investor Literacy Quiz to identify your areas of strength and weakness...then use our tutorials to fill in your weaker areas

3. We create test vignettes, in this particular experiment creating 60 different combinations for each respondent. The combinations, called vignettes or concepts in other published work, comprise 2-4 elements, each element appearing five times. The elements appear against different backgrounds since all the elements vary from one vignette to another. The underlying experimental design, a *recipe book* controls which particular elements appear in each vignette. Although to the untutored eye the 60 different vignettes appear to be simply a random, haphazard collection of elements with no real structure, nothing could be further from the truth. The experimental design is a well-thought out mathematical structure ensuring that each element appears independently of every other element, repeated the same number of times across each element., This allows us to deconstruct the response to the 60 test vignettes into the individual contribution of each element. Statistical analysis by OLS, ordinary least-squares regression, will immediately reveal to us which elements are responsible for the rating and which simply go along, not contributing anything.

*One final note is important. In most of these experimental design studies the researcher selects one basic design and populates that design once, in our case populating the 60 vignettes with the relevant elements called for by the design. The next step in conventional research would present the SAME design to each new respondent, perhaps randomizing the order of the elements. Thus, if we were to work with 200 respondents in the experiment, each respondent would be exposed to the same set of 60 vignettes so that in the end each vignette would be generate 200 ratings. The order of the vignettes in*

*I approach here differs dramatically. Each respondent evaluates a unique, new set of combinations following the same basic experimental design, but the design is permuted, changed a bit. Each respondent ends up evaluating the same set of elements the same number of times, but the permutation changes the particular combinations. In effect, each respondent's experiment covers new territory in the space defined by the elements. The result is a robust estimate of the contribution of each element when it comes time to deconstruct the responses into the combinations of the components.*

4. We see an example of a vignette in Fig.2. The bottom of the vignette shows the rating assigned to the vignette. The respondent reads the vignette in its entirety and rates the vignette on the scale below the vignette. Typically, the experiment is run on a computer attached to the Internet. A program sets up the vignettes locally on the respondent's computer, presenting the vignettes, combinations of elements according to the experimental design, and acquires the rating. The interview is relatively quick, requiring about 15 minutes for the presentation of the vignettes followed by

a short classification questionnaire. Several million respondents have participated in literally thousands of studies. As long as the respondent is at least a bit interested and participates, the field execution of the study with respondents is straightforward. The process is automatic from the start of the experiment to the data analysis, making the system scalable. The experiment is designed to create a corpus of knowledge in many different areas, ranging from marketing to food to the law, education, and government

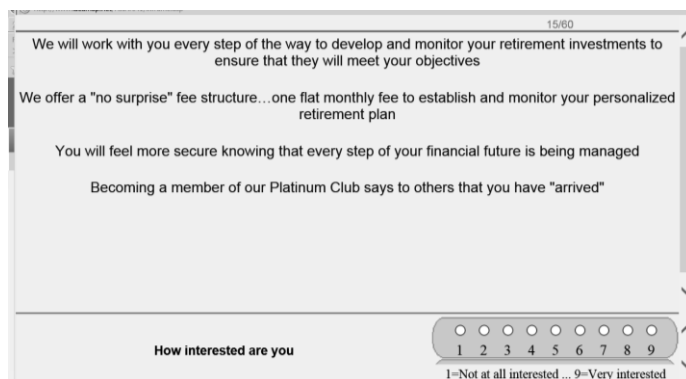


Fig.2: An example of a test vignette. The elements appear in centered rectangular format with no effort to connect the elements, a format which enhances 'information grazing.' The vignette shows the ratings scale at the bottom, and the progress in the experiment at the top right (screen 15 out of 60).

5. The data from each respondent allows us to build an individual level model or equation. Building the model at the level of the individual is the most powerful format of control, known to statisticians as the strategy of 'within subjects design.' The experimental design allows OLS (ordinary least-squares) regression to create the model, then equation, relating the presence/absence of the 36 elements to the rating assigned by the respondent. The original rating scale that we see at the bottom of Fig.2 is known as a Likert scale, or category scale, an ordered set of categories representing the psychological range from 1 (not at all interested) to 9 (very interested). For our analysis we simplify the results, focusing on two parts of this 9-point scale, with the lower part (ratings 1-6) corresponding to not interested and the upper part (ratings 7-9) corresponding to interested. We recode ratings of 1-6 to the number 0 and ratings of 7-9 to the number 100.

*Although the 9-point scale provides more granular information, the reality is that managers focus only on one thing: will the prospect be interested in becoming a client of the services. The 9-point scale is more useful when it is turned into a binary scale.*

*Note: Often when the notion of binary scale is introduced, based on the proclivities and interest of the manager, the question is raised as to why not simply have the respondent in this micro-science rate each vignette on two points, not interested versus interested. That strategy is certainly possible, but there are some analyses, specifically*

clustering, which perform a bit better using the more granular 9-point ratings. And so we keep both, reporting our results using the binary model (also called the INT Model), but doing the cluster analysis specifically using the 9-point model (also called the PER Model).

6. Some of the particulars underlying the modeling are:
  - a. The models are created at the level of the individual respondent using the well-accepted procedure of OLS, ordinary least squares regression.
  - b. The experimental design ensures that the 36 elements are statistically independent of each other so that the coefficients, the impact values of the elements, have absolute value. The inputs are 0/1, 0 when the element is absent from a vignette, 1 when the elements is present in the vignette.
  - c. OLS work at the individual respondent level, creating first the PER Model (dependent variable is 1-9 from the rating scale), and then the INT Model (dependent variable is 0/100, 0 for ratings of 1-6 on the 9-point rating scale, 100 for ratings of 7-9 on the 9-point rating scale).
  - d. OLS uses the 60 sets of elements/ratings, one per vignette, as the cases. There are 36 independent variables and 60 cases, allowing sufficient degrees of freedom for OLS to emerge with robust estimates
  - e. We write the equation or model as:  $\text{Rating} = k_0 + k_1(A1) + k_2(A2) \dots k_{36}(D9)$ .
  - f. The equation says that the rating is the combination of an additive constant,  $k_0$ , and weights on the elements. The elements appear either as 0 (absent) or as 1 (present), so the weights,  $k_1 - k_{36}$ , showing the driving force of the different elements.

C. What the data tell us (Table 2)

As noted above, we create an individual model for each respondent, which we can do because the 60 vignettes tested by a respondent constitute an experimental design. We average the corresponding parameters of all relevant respondents for a subgroup in order to estimate the average parameter for each subgroup.

We now look at the strongest performing elements from the average INT Model, that model which related the presence/absence of the 36 elements to the binary rating (1-6 transformed to 0, 7-9 transformed to 100). The strongest performing elements appear in Table 2. The table shows all elements which generate an impact value or coefficient 8 or higher for any key subgroup, whether total sample, gender, age, or income, respectively.

Beginning with total sample and then moving across to the other groups, we interpret the results as follows:

1. The total panel comprises 241 respondents. We can break out the groups into gender, age, and income.

Table 2: Strong performing elements for the Total Sample and for key subgroups defined by how the respondent classifies himself. The table presents only those strong-performing elements with average impacts of 8 or higher from the INT models

	Total Sample	Male	Female	Under 40	Over 40	Low Income	High Income
Base size	241	44	197	107	134	199	42
Additive constant	35	28	36	29	40	34	37
Tell us when you want to retire and we will develop an action plan to get you there	7	9	6	9	6	7	6
Assess your financial health with our retirement planning worksheet...answer some basic questions about your financial present and your goals for the future...then we will recommend some strategies to get you there	6	6	6	9	4	8	0
Convenient 24/7 online access to your account and our experts...ask questions...monitor your progress...all online!	6	8	5	8	5	7	0
We will work with you every step of the way to develop and monitor your retirement investments to ensure that they will meet your objectives	5	8	4	6	4	5	3
Planning for your retirement now means you won't get caught short since you still have time to do something about it	3	8	2	2	5	3	4
It is never too early to start planning for your financial future...the steps you take today will significantly affect the quality of your life 10,20 even 40 years from now	3	8	2	5	2	2	9

2. The additive constant tells us the conditional probability of a person saying *interested in* what the financial advisor has to say, i.e., assigns a rating of 7-9, when reading a vignette which has no elements (the baseline). Of course all vignettes comprised elements, so the additive constant is an *estimated* parameter. We can use the additive constant as a baseline. For the total panel it is 35, meaning that 35% of the respondents would rate a vignette 7-9. Males are less likely to be positive while females are more likely to be positive (additive constants of 28 vs. 36). Those under 40 are far less likely to be positive, those over 40 are more likely to be

positive (additive constant of 29 vs. 40). Income makes no difference.

3. Beyond the baseline are the elements, which contribute to the total. We add up to four elements to the baseline to get an estimated total value, i.e., the percent of respondents who say that they would be interested in the vignette about the financial consult were the elements to be part of the advertising.
4. Our elements, by and large, are low. Table 2 shows the strongest elements only, and only elements which generate a coefficient or impact value of +8 for any one of the subgroups. We interpret that +8 to mean that when the element is incorporated into the advertising vignette at least 8% more people will rate the vignette 7-9, i.e., say 'I'm interested.'
5. Our first look into the results suggests nothing particularly strong emerges from the total sample. We do see six elements scoring well in at least one subgroup. However, we see no general pattern. That is, we don't see an element working very well across the different groups. Furthermore, reading the different elements only confuses us. There are no simple patterns.
6. Our first conclusion, therefore, is that the experiment worked. We are able to develop elements, test combinations, deconstruct the combinations, and identify winning elements. The experiment, at least thus far, does not reveal to us deeper information about the *mind* of the prospect.

#### *D. Deeper, possibly more fundamental structures of the mind by clustering (Table 3)*

Up to now we have looked at people as individuals, perhaps falling into convenient groups defined by easy-to-measure variables such as gender, age, income. We could multiply our easy-to-measure variables by asking our respondents lots of questions about themselves, about their attitudes towards financial investors, about their feelings towards risk versus safety, and so forth. And we could then classify the respondents by the different groups to which they belong, searching for a possible co-variation between group membership and response pattern to elements.

The just-described approach typifies the conventional way of thinking about people. We define people as belonging to groups and then search out the linkage between such groups and some defined behavior. Scientists call this strategy the hypothetico-deductive method, beginning first with a sense of 'how the world might work,' and then running an experiment to confirm, or just as likely to falsify, that hypothesis. We work from the top down, thinking about what might happen and proceeding merrily to validate or reject that thinking.

Let's proceed in a different manner, without hypothesizing about how the world works. Let's proceed with the data we have, looking instead for basic groups who show radically different, interpretable patterns. In the world of color this is analogous to looking for the basic colors of the spectrum, red, yellow, blue, which must emerge out of the measured thousands of colors of flowers. Let's work from the bottom up, in a more pointillist, empirical fashion, emulating Francis Bacon in his *Novum Organum*.

How then do we do this? How do we find naturally occurring groups of people in a specific population who show different patterns of behavior or at least responses for the micro, limited area? That is, we are working with a small corner of reality, one's responses to messages about choosing a financial advisor. It's a limited aspect of reality. How is that reality constituted? Are there different groups of minds out there wanting different features? Are these groups of minds interpretable? To continue with the aforementioned metaphor, can we find the *basic colors* for this aspect of reality, the red/blue/yellow, not of the whole world, but the red/blue/yellow of choosing a financial advisor?

That we have limited our focus to the limited, micro area of messaging for client acquisition by a financial advisor makes our job easier:

1. We are working in a corner, nook, a little region of *reality*.
2. We already have rich material produced by our study. Our study with 36 elements and 241 profiles of impact values tells us how 241 individuals value the individual elements.
3. So with that small wedge of reality, let us see whether there is a deeper structure, focusing only on the reality of choosing a financial advisor and using only the *mind* of the consumer as a way to organize reality.
4. Continuing our metaphor of colors, we have come upon a new limited aspect of reality. What are the basic dimensions of that new, limited aspect of reality?
5. We have only two ground rules.
6. Ground Rule 1: We should be looking for primaries, the fewer the better, for this new aspect of reality, our *mind of selecting the investment advisor*.
7. Ground Rule 2: We must be able to interpret these primaries, in a simple way. They must make sense.

The foregoing introduction leads us naturally to our data, our 241 rows (one per respondent), and our 36 columns (one per element). Whether we use the INT Model (binary 0/100 value as the dependent variable) or the original PER Model (1-9 rating as dependent variable) makes little real difference. For clustering, leading to viewpoints, clusters, segments, we will use the original PER Model. There is more granularity in the 9-point scale than in the binary INT scale.

Our analysis is straightforward. We apply the method of cluster analysis to our 241 rows x 36 columns (respondents by

coefficients, PER Model). We do not incorporate the additive constant into our cluster analysis, because it doesn't give us information about the response to particular elements, the focus of the cluster analysis.

1. Cluster analysis puts our 241 respondents first into two groups, then into three groups, then into four groups, and so forth. These are clusters, which we can call *viewpoints* because they represent different viewpoints that people have about what is important in the interaction with a financial advisor. Furthermore, the word 'viewpoint' emphasizes the psychological nature of the cluster, that we are dealing with the *mind* here, the mind as it organizes one small corner of reality, the approaching interaction with a financial advisor.
2. We look at the average values from our INT Model (binary transform) to see what elements float to the top, i.e., score well, for each viewpoint.
3. We end up with a solution suggesting three different viewpoints, as Table 3 shows us. These three viewpoints are shown and named by virtue of the strongest performing elements in each viewpoint.
4. And we end up with one viewpoint which is our target (Viewpoint 3), and two groups that are not targets, and might for actual implementation be put together (Viewpoints 1 and 2).

Let's visit the data to see what cluster analysis has delivered to us, keeping in mind that we used the PER Model to get to the clusters, but we look at the data expressed in the binary, INT Model. It's easier to understand the INT data which talks to us about a person being part of a group or not part of a group. The data appear in Table 3.

1. We end up with seven elements scoring well, which we defined operationally as having a coefficient or impact value in the INT model of 8 or higher for any key subgroup.
2. We present three sets of columns, the first for the total sample, the second for Target (Viewpoint 3) vs. Non Target (Viewpoints 1 and 2 combined), and the third for all three viewpoints separately.
3. The additive constants, our baselines, are all 30-40. There is no Viewpoint Group just ready to spring to attention, willing to buy the services of the financial advisor.
4. The real differences come from the elements as responded to by the individuals in the different Viewpoint Groups.
5. The total sample shows no strong elements. This means that without any knowledge of the mind of the prospect it's unlikely that someone will know what to say, or the right thing to say. Perhaps the strongest message, with a coefficient of +7 (additional 7% interested in working with the advisor) is the phrase: *Tell us when you want to retire and we will develop an action plan to get you there.*
6. Our most promising group is Viewpoint 3, comprising 70 of our 241 respondents, or 28%. Identify those individuals and there are combinations of messages that allow us to get 80% of the prospects interested in what we have to offer.

7. On the other hand, the complementary viewpoint, NON-PROSPECTS, comprising the remaining 171 respondents, does not show any strong elements at all. Individuals in the NON-PROSPECT viewpoint are simply not interested in what the financial advisor has to say.
8. Looking to the far right, the last three columns, we see the differences in the performance of these strong elements across the three segments (V3 TARGET: Up to date full service retirement planning; V2 You know Me and Can Help; V1 Easy Retirement Planning.)
9. It will be our job to to we assign a new person to a Viewpoint Group. Table 3 simply tells us what to say, precisely, once we find the people, a major advance over knowledge that we began with, but not the whole story.

Table 3: Performance of the very strongest elements in the three Viewpoints (mind-set segments) emerging from the cluster analysis. People in Viewpoint Group V3 become the target group to be identified as the promising clients for the financial advisor.

	Total	V3	V1 + V2	V2	V1
Base size	241	70	171	100	71
Additive constant	35	32	38	33	40
Tell us when you want to retire and we will develop an action plan to get you there	7	17	3	0	8
It is never too early to start planning for your financial future...the steps you take today will significantly affect the quality of your life 10,20 even 40 years from now	3	12	0	2	-3
Your plan will be reviewed annually not just by YOUR consultant but with our entire staff of retirement planning experts...working in collaboration to ensure that your needs are being met	4	11	1	0	3
Convenient 24/7 online access to your account and our experts...ask questions...monitor your progress...all online!	6	11	4	3	6
During your peak earning years we will help you to manage the changes in your life to the advantage of your longer term retirement plan	1	11	-2	0	-5
We offer a "no surprise" fee structure...one flat monthly fee to establish and monitor your personalized retirement plan	1	11	-3	3	-13

*E. Finding Viewpoints (Minds) in a population of bodies, nodes (objects which behave)*

We walk around with lots of numbers attached to us. The competent data scientists can extract information about us from our tracks, whether these tracks be left by our behavior on the Web (sites that we have seen), by forms that we have filled out and are commercially purchasable (e.g., through Experian or Trans Union or any of the other commercial data providers, by loyalty programs), or even by questionnaires that we ask respondents.

All of the available data, properly mined, collated, analyzed, and reported, might well tell us when a person is ready to hire a financial advisor, e.g., upon the occasion of marriage, a child, a promotion, a job change, a move to another city, and so forth. But just what do we say to this particular prospect, the person standing before us in person, or interacting with our Web site, or even sitting at home destined to be sent a semi-impersonal phone message, email, or letter? What are the precise words to say to this person?

Those in sales know that a good sales person can sense what to say, that 5, 10, 20 years of experiences hones one’s perception of just who this prospect is, and what will sell. But how do we scale this knowledge so we know what specifically to say to a specific person, perhaps a person we don’t even know, haven’t even met, might never meet? Or restated, how do substantially increase the probability of sending the right message to the right person at the right time?

The foregoing results in Table 3 show us what to say and to whom. The problem now becomes one of determining the viewpoint to which a specific person belongs. Unfortunately, people do not come with brass plates across their foreheads telling us the viewpoints to which that person belongs. And there are so many viewpoints to discover for a person, as many sets of viewpoints as there are topic areas for Mind Genomics®. The bottom line here is that data scientists working with so-called Big Data might be able to infer that a person is likely to be ready for a financial advisor, but as currently constituted, the same Big Data is unlikely to reveal the Viewpoint Group to which the individual person belongs. We have petabytes of data, reams of insights, but not the knowledge specificity about the way the *mind* works for any specific, operationally defined topic in the reality of our experience.

We move now to the second phase of our work reported here, discovering the viewpoint to which any person belongs. We have already established the micro-science for the financial planner, the set of phrases to use, the different Viewpoint Groups. We know from our 241 respondents the Viewpoint Group to which each person belongs, having established group membership by used cluster analysis. How then do we identify any new person, anywhere, as belonging to one of our three Viewpoint Groups, and thus know just what to say to that person?

We follow the steps below which represent one systematic way, almost algorithmic in nature, to move from the micro-science, the ‘what’ of the reality, how that reality is constituted, to the ‘who,’ fitting a new person into that reality.

1. Our first known is the Viewpoint Group membership for each of our 241 respondents.
2. Our second known is the individual PER Model for each person, the set of 36 coefficients for the 36 elements, along with the additive constant. The PER Model allows us to estimate the likely 9-point rating for each element for each respondent. With the PER Model, we can also estimate how each element would score on a 1-3 scale. We will use the latter transformation, working with an easy-to-use 3-point scale.
3. Thus, Step 1 gives us Viewpoint Group, and Step 2 gives us the expected rating of each of the 36 elements by each of the 241 respondents, on a 1-3 scale.
4. We use discriminant function analysis (DFA), similar to regression analysis, to identify which of our 36 elements best discriminates across the three Viewpoint Groups. Table 4 tells us that there are three elements which discriminate most strongly across the three Viewpoint Groups. Table 4 does not tell us how to use these discriminating elements to score new people as belonging to a Viewpoint Group.

Table 4: The three most discriminating elements with respect to how the respondents fall into the three Viewpoint Groups. The element text is on the left. The ability to discriminate across the three segments (F Ratio) appears on the right.

Element	F Ratio
Becoming a member of our Platinum Club says to others that you have “arrived”	62.98
It is never too early to start planning for your financial future...the steps you take today will significantly affect the quality of your life 10,20 even 40 years from now	14.13
Convenient 24/7 online access to your account and our experts...ask questions...monitor your progress...all online!	10.43

5. We develop a simple three-point questionnaire that when used by a respondent or by the sales person fits the prospects into one of the three segments. We embed this questionnaire into a simple cloud-compatible Web app, shown in Figures 2-8. When we use the questionnaire to predict Viewpoint Group membership for three viewpoint groups, we end up assigning the right person to the right Viewpoint Group 63% of the time. When we use the questionnaire to predict Viewpoint Group membership for two groups, prospects versus non prospects, we end up assigning the right person to the right Viewpoint Group 78% of the time.
6. Fig.2 shows the first page, the landing page. This page tells the respondent what to do. The screen is written from the point of view of a sales person talking to the prospect, with the sales person typing in the answer.



On the following screens, we're going to give you 3 simple questions to ask YOUR customers.

Ask each question, type in the answer...that's all you have to do.

Then... the CUSTOMER ENHANCEMENT TOOL will suggest what you should FOCUS ON..and of course..what you should AVOID!

>>

Fig.2: The viewpoint identifier: Orientation page

7. Fig.3 shows the three questions. These elements come from the DFA, the aforementioned discriminant function analysis. The questions appear in a randomized order. Often there are dummy questions inserted into the set of three questions. The dummy questions prevent the user from *gaming* the system.

Overall, how interested are you in this retirement plan if...  
Our Retirement Resource Center has the tools, information and experts to help you make smart retirement planning decisions

Not at all      Not sure      Definitely

Overall, how interested are you in this retirement plan if...  
Becoming a member of our Platinum Club says to others that you have "arrived"

Not at all      Not sure      Definitely

Overall, how interested are you in this retirement plan if...  
We offer a "no surprise" fee structure... one flat monthly fee to establish and monitor your personalized retirement plan?

Not at all      Not sure      Definitely

Fig.3: The three questions in randomized order, used to discover the Viewpoint Group of a prospect.

8. Once the respondent or the salesperson completes the question, the system determines the respondent's Viewpoint Group, returning with the proper messages to use, and just as important, the messages to avoid. Fig.4 shows what to see for respondents assigned to the target Viewpoint Group, V3. This Viewpoint Group can be described as wanting 'Up to date full service retirement planning.' Prospects belonging to Viewpoint Group V3 will end up being our best prospects because they are the only Viewpoint Group showing very strong responses to elements.

Addressable Mind Genomics Survey

Welcome to the CUSTOMER ENHANCEMENT TOOL™

**Up to date full service retirement planning**

Things to say

- Tell us when you want to retire and we will develop an action plan to get you there
- It is never too early to start planning for your financial future... the steps you take today will significantly affect the quality of your life 10,20 even 40 years from now
- We offer a "no surprise" fee structure... one flat monthly fee to establish and monitor your personalized retirement plan
- During your peak earning years we will help you to manage the changes in your life to the advantage of your longer term retirement plan

Things to avoid

- We offer a combination fee structure... a low monthly fee plus discounted prices for each transaction
- Available exclusively to our VIP customers, a complimentary consultation with someone from the Wills and Trusts department...ask the questions you have been meaning to ask
- A special class of service is available to those who attain Platinum status
- Becoming a member of our Platinum Club says to others that you have "arrived"

Fig.4: What to say to respondents who are assigned to Viewpoint Group V3, the group most likely to purchase services from the financial advisor. This segment has been labeled 'Up to date full service retirement planning' based upon the elements to which these individuals are most responsive.

9. Fig.5 shows what to say for prospects assigned to Viewpoint Group V2, which we label 'You know Me and Can Help'. This Viewpoint Group is far less responsive to the messages and thus, not part of the easiest-to-sell prospects for the financial advisor.

Addressable Mind Genomics Survey

Welcome to the CUSTOMER ENHANCEMENT TOOL™

**You know Me and Can Help**

Things to say

- Assess your financial health with our retirement planning worksheet... answer some basic questions about your financial present and your goals for the future...then we will recommend some strategies to get you there
- Take our Investor Literacy Quiz to identify your areas of strength and weakness...then use our tutorials to fill in your weaker areas
- We will work with you every step of the way to develop and monitor your retirement investments to ensure that they will meet your objectives
- We offer you pointers, recommended reading and worksheets so you can develop a retirement plan that suits your needs

Things to avoid

- As you near retirement... you can take steps to improve how long and how well your hard-earned savings will work for you during your retirement years
- Our philosophy is based on interpersonal relationships so we will want to meet with you and get to know you...this enables us to develop a better retirement plan for YOU
- Our Retirement Resource Center has the tools, information and experts to help you make smart retirement planning decisions
- You will feel more secure knowing that every step of your financial future is being managed

New

Fig.5: What to say to respondents who are assigned to Viewpoint Group V2 who are less likely to purchase the services of the financial planner.

10. Fig.6 shows what to say to prospects assigned to Viewpoint Group V1, Easy Retirement Planning. This Viewpoint Group is far less responsive to the messages and thus, not part of the prospects.

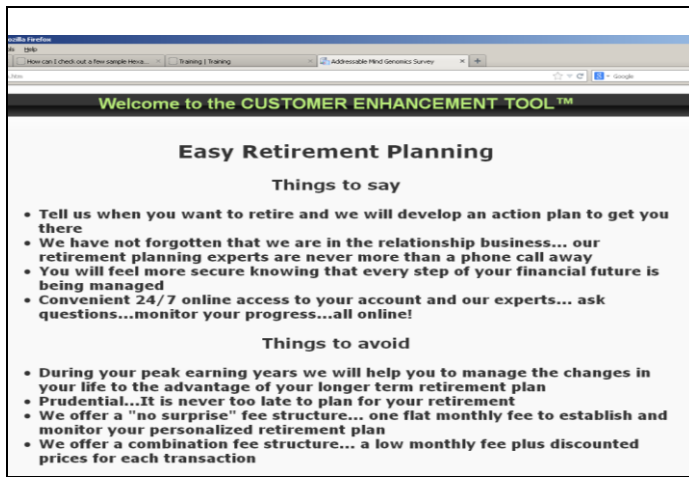


Fig.6: What to say to prospects who are assigned to Viewpoint Group V1, who are less likely to purchase the services of the financial planner.

11. We combine Viewpoint Groups 2 and 1 into non-prospects, and repeat the foregoing approach. This time we end up with two results, prospects (Fig.7) and non-prospects (Fig.8). Note that Fig.7 is the same as Fig.4, both referring to Viewpoint Group V3.

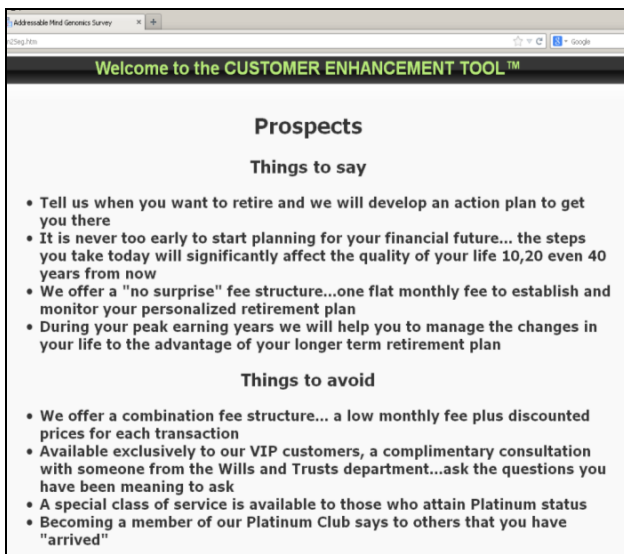


Fig.7: What to say to respondents who are assigned to Prospects (the same Viewpoint Group V3), the individuals most likely to purchase services from the financial advisor. This segment has been labeled 'Up to date full service retirement planning' based upon the elements to which these individuals are most responsive.

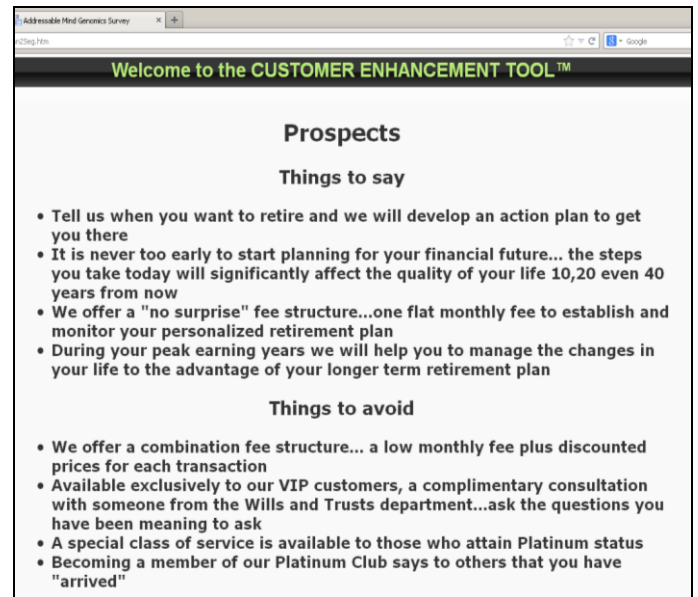


Fig.8: What to say to respondents who are assigned to Non-Prospects (the larger group, comprising respondents in the original Viewpoint Groups 2 and 1). These respondents are not likely to buy from the financial advisor, at least based upon any of the 36 elements tested.

12. As a demonstration, the two Web sites below provide tools to assign a new prospect either to one of the three Viewpoint Groups, or to the Prospects/Non- Prospects groups.

3 Viewpoint groups:

<http://www.mjiweb.com/mjitt/FinPlan/FinPlan.htm>

Prospects vs. Non-Prospects:

<http://www.mjiweb.com/mjitt/FinPlan2Seg/FinPlan2Seg.htm>

#### IV. REVISING BIG DATA WITH VIEWPOINTS TO CREATE BIG MIND + BIG DATA

Up to now we have been dealing with small groups of individuals whose specific mind-set or viewpoint in a specific, limited topic area we can discover, and then act upon. But what happens when we want to deal with thousands, millions, and even billions of new people? Consider, for example, the points in Fig.9, top panel. Consider these points as individuals. We know from behavior how they are connected with each other. There are many visualization techniques which create the interconnections based upon one or another criterion. And from these visualizations we can ascribe something to the network. We haven't shown the connections among the points. The corrections, the edges, would make confusing pictures.

What happens, however, when we know the mind of each person, or at least the membership in, say, 12 different topic areas? What power would emerge to understanding behavior in a profound way if we were to know both the network itself, what points link together by specified common behaviors, e.g., shopping, AND the mind of each point, each node in these

12, or 120, or 1000 defined topic areas? The bottom panel of Fig.9 shows what happens when we enrich our knowledge of the network by the underlying pattern of the *mind* behind the point, the node in the graph. We now know the network *and the mind* of these connected nodes. We can create a web of interconnected points and discover some of the commonalities of the points, not based on who the points are or what the points did, but rather how the points think about many relevant topics.

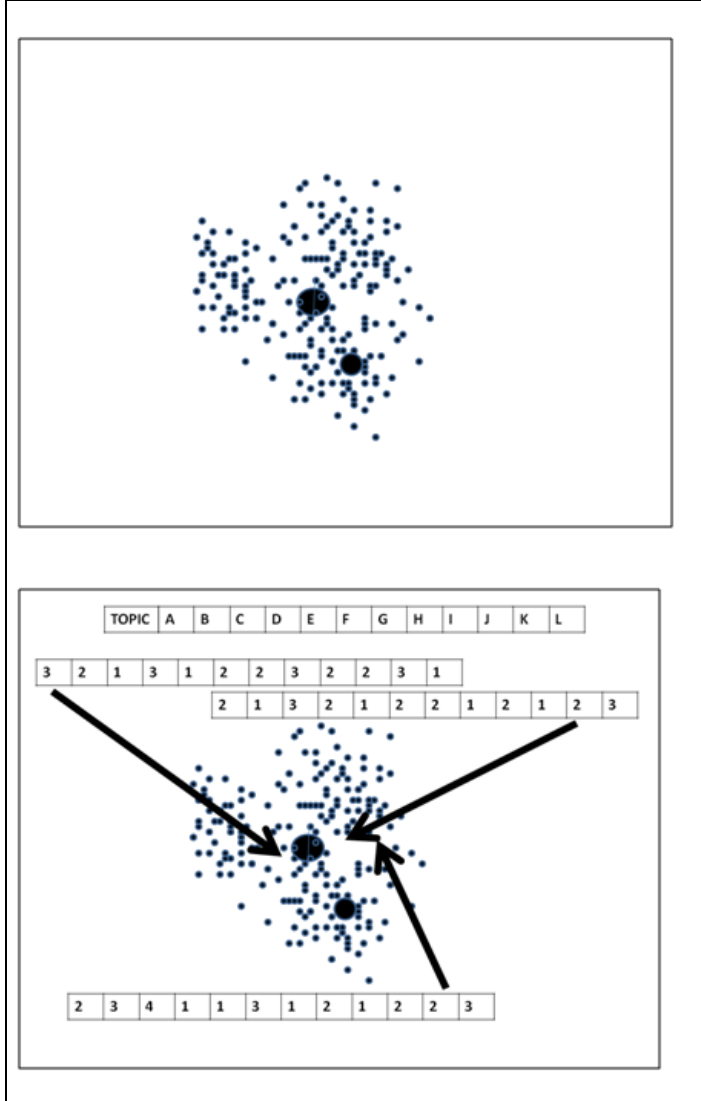


Fig.9: Example of nodes (i.e., people), perhaps connected by a network. The top panel shows the network of people as points. The bottom panel shows the potential of knowing the mind of each person, i.e., each point in the network.

### B. A new vision of Big Data – Big Data and Big Mind

How do we move from Mind Genomics of one topic, say our choice of financial advisor, to many topics and then through typing thousands, not hundreds, and finally millions, tens of millions, hundreds of millions, or even billions? And how do we create this new web, this New Mind + Big Data in a

way that's empirically sound, extensive, actionable, and archival for decades? In essence, how do we go from a map of nodes to a map of connected minds?

To reiterate, our goal is to understand the specific Viewpoint Group memberships of each point on the network, where the point corresponds to a person. The big picture is thus millions, perhaps hundreds of millions of points, people, observed two ways:

1. Their interactions with each other, as measured objectively, either by who they are or by how they behave, such as what they view on the Web, what they order, with whom they interact in conversations.
2. Their minds, or at least the pattern of membership in Viewpoint Groups as determined through Mind Genomics, for a specific set of different topic areas. There may be as few as one topic area, or several dozen or even 100 or more.

We follow these steps:

1. Identify the different topic areas. In Fig.9 we show a vector of 12 topic areas. In practice the number of topic areas is unlimited. Our procedure should be able to score each point, person, node, on every topic area.
2. Create the 12 Mind Genomics Micro-Sciences, one Micro-Science for each of the 12 topics area. We follow the same steps that we followed above to understand the messages for selecting the financial advisor. If we want to study a wider array of topic areas, we do so easily. Just note that each topic area is studied separately, working with 200-300 relevant respondents.
3. Identify the relevant Viewpoint Groups for each topic area, just as we did for the financial advisor. We analyze the Micro-Sciences separately, with different and appropriate respondents for each topic area.
4. Create the Viewpoint Identifier for each topic area, using DFA, and the data we collected. We saw the typing tool for one topic area in Figures 2-8. Those figures provide us with a template.
5. Invite 10,000 people to participate in the typing, by direct interaction, using the Web-based tool similar to that in Fig.3, but particularized to the topic area.
6. From the 10,000 people invited to participate for a topic area, perhaps 1000 will respond. For each person who responds determine the Viewpoint Group to which that person belongs for the topic area. Find out additional information about these individuals from either purchasable data about them (third party), or ask them to fill out a detailed self-profiling classification.

7. Step 6 provides us with new material on which to do a second DFA for the topic area. For each person we know both the Viewpoint Group membership, as well as have a vector of purchasable information about that person. Run the DFA once again, this time predicting the person's Viewpoint Group membership (provided by the interaction) to the purchasable information about the person. The predictability will be somewhat attenuated because the DFA will work on relating Viewpoint Group Membership to exogenous information about the purchase, not to the response to elements, which is about the person's mind.
8. Step 7 provides us a scoring tool, from the DFA, allowing us to go through the database, small or large, and score each person in terms of Viewpoint Group membership based upon the classification function developed by the DFA for that particular topic.
9. Repeat the steps above for each topic area creating a separate database scoring tool for each topic area.
10. We now have the tool by which to predict the Viewpoint Group membership of each person, topic by topic, as long as we are given the relevant external information or self-profiling classification information.
11. The foregoing steps end up turning so-called Big Data into Big Mind + Big Data. By way of closing, when we have 100 topics, not just 12, each person can be assigned to the relevant Viewpoint Group for each topic area, following the previous systematic, sequential approach.

#### ACKNOWLEDGMENT

The authors thank Linda Ettinger Lieberman, co-author and Editorial Coordinator, Moskowitz Jacobs, Inc., for preparing this manuscript for publication.

#### REFERENCES

- [1] C. Jung, *Psychological Types: A Revision*. (The Collected Works of C.G. Jung). Princeton, NJ: Princeton University Press, 1976.
- [2] A. Gofman and H.R. Moskowitz. "Application of isomorphic permuted experimental designs in conjoint analysis," in *Journal of Sensory Science* 25(1), pp. 127-145 (19), 2010.
- [3] H. Moskowitz, B. Batalvi, and L. Ettinger Lieberman, *Empathy and Experiment: Applying Consumer Science to Whole Grains as Foods*. Proceedings of the Whole Grains Summit 2012. Minneapolis, MN: AACC International, Inc., 2013.
- [4] H. Moskowitz and A. Gofman. Rule Developing Experimentation: A Systematic Approach to Understand & Engineer the Consumer Mind. Bentham Science Publishing, 2012.