

Review Article

Advancing Mind Genomics: Using AI (Artificial Intelligence) to Generate Topics, Questions, Messages as Answers, and 'Synthesized' Minds (Respondents)

Howard R Moskowitz^{1*}, Stephen Rappaport², Tim Morford¹, Martin Mulvey¹, Kenneth Tomaro¹, Sunaina Saharan³ and Angelica DiLorenzo³

¹Cognitive Behavior Insights, Inc., Albany, New York, USA

²SDR Consulting, Norwalk, CT, USA

³Global Population Healthcare Management Forum, Brooklyn, NY, USA

*Corresponding author: Howard R Moskowitz, Cognitive Behavior Insight, Inc., Albany, New York, USA

Received: December 01, 2023; Accepted: December 07, 2023; Published: December 15, 2023

Abstract

AI created a resource bank of statements about what a doctor might say to a child in order to deal with the child's obesity. After the AI generated messages were developed, 16 of the messages (elements) were selected and combined into vignettes according to an underlying experimental design, whose specific combinations differed for each 'respondent'. Each set of 24 vignettes comprised a stand-alone set of combinations and were evaluated by AI prompted to act as a specific person in the medical world (receptionist, doctor with 10 years of experience, nurse practitioner). Deconstruction of the ratings by regression showed the contribution of each AI created message to the rating scale. The coefficients 'made sense' when the regression was done according to 'WHO' the AI was defined to be. Further clustering the coefficients across the respondents revealed two clearly different mind-sets. The systematic approach using AI as both the provider of ideas and the evaluator of these ideas presents a new vista for learning about how to communicate with people, using technology to dramatically accelerate and fine-tune that learning.

Introduction

It is hard to overestimate the excitement with which AI, artificial intelligence, has been greeted and adopted, especially since the introduction of Chat GPT, and associated technology [1]. One can scarcely read of any topic of human endeavor without one or another pundit bringing up the impact of AI for that field. This paper looks at the potential of AI to synthesize 'respondents', with the goal of accelerating the learning of professionals who want to learn to counsel people in nutritional health. The use of synthesized people, personas, is not new, and has been a topic of interest for some decades now [2-4]. What is new is the vision of moving beyond personas derived from large scale studies to one-off studies created entirely by synthetic means.

This paper combines AI with Mind Genomics to develop a new system for training and education. The underlying vision is to accelerate knowledge development by having the AI provide topic-relevant ideas (Idea Coach in Mind Genomics), and then 'personas' created by artificial intelligence, based upon combinations of features of the way people think, who the people are etc.. The raw material for these personas come from 'self-profiling classification questions' that a researcher might ask a human respondent. In short, the paper presents a Mind Genomics study, with the materials, from beginning end under the control of the machine, not the researcher. All the researcher does is select ideas at the very start of the study, these ideas later being tested in the study itself.

The process of Mind Genomics begins with a series of questions, those questions telling a story, and then for each question provide four different answers. The actual experiment consists of combining these answers into small combinations of 2-4 answers, at most one answer from a question, but often no answer from a question. The 24 combinations, called vignettes, are evaluated by respondents. Each respondent evaluates a set of 24 different vignettes, the uniqueness of the sets of 24 vignettes ensured by a permutation algorithm [5]. The analysis, done by OLS, ordinary least squares regression, shows through the coefficients of a linear equation the degree to which the 16 elements, viz., answers to the question, drive the respondent's rating.

During the evolution of Mind Genomics, a process taking place for more than 30 years, since its introduction in 1993 [6-8], the consistently hardest part of any effort, basic research or applied research, was developing the questions, and to a far lesser degree coming up with the answers to the questions. More than one study ended up being aborted simply because the researcher could not generate the four questions which tell the story. Once the four questions were developed, for the most part, researchers were able to push through to the four answer for that question. A solution to the issue of frustration in the creation of questions and answers emerged with the incorporation of AI into the process.

A third problem, motivating people, respondents, to participate, proved to be simple to solve because of the emergence of companies which provided paid panelists. Money solved the problem of

participation, if not motivation. This paper will deal with the introduction of synthesized respondents, to reduce cost, and to facilitate new types of systematized investigations not possible before.

Mind Genomics requires systematized thinking about a topic, an ability that all too often needs to be learned through coaching. It was to provide this coaching that the original AI was introduced in mind-year, 2023, in the form of Idea Coach. This paper focuses on the further introduction of AI 'synthesized respondents', in an effort to make the Mind Genomics process a streamlined one, from beginning to end, appropriate for teaching as well as for practical application.

Explicating the Process-Part 1-Setting Up the Study Using AI

Figure 1 (Panel A) shows the request by the Mind Genomics set-up 'template' for the four questions, each to be addressed by four answers. It is the development of these four questions which become a difficult hurdle. The creation of the Idea Coach allows this topic to be addressed. Figure 1 (Panel B) shows the rectangle where the researcher can write out the question. The question posed to AI through Idea Coach is very simple: *I am a doctor treating obesity in children. How do*

I talk to the children to make them understand. One could further fine-tune the Idea Coach by telling it to explicate the question in discussion form, as well as to provide questions of length 15 words or less, and questions understandable to a 12-year-old. Those statements become part of the query. The actual process is made as simple as possible so that the effort focuses on the topic.

The actual short description appears as the topic in Table 1. The Idea Coach was run three times to generate a reservoir of questions. The Idea Coach is typically run 5-10 times, providing information which ends up teaching the researcher. The Idea Coach is also run several times for each of the four questions. The combination of different sets of 15 questions for a topic description and different sets of 15 answers for each question provides a unique resource booklet on the topic. Each iteration in Idea Coach lasts about 20 seconds, so that in five minutes one can produce 15 sets of 15 questions each. In the end, only four questions will be chosen.

The results from the first 15 questions appear in Table 1. The questions emerging from the AI embedded in Idea Coach are easily understood by a human being.

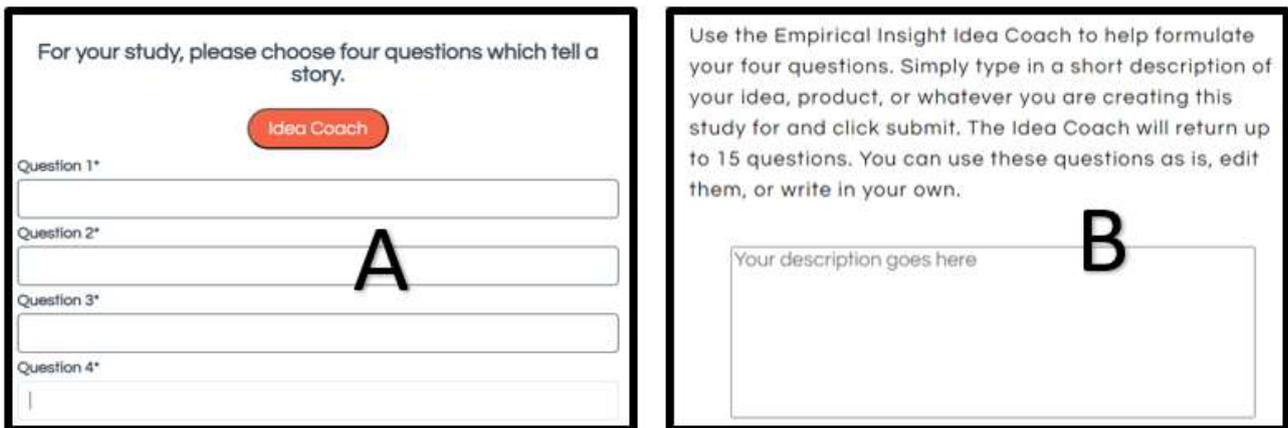


Figure 1: The templated screens requesting the researcher to provide four questions. Panel A shows the request for four questions which tell a story Panel B shows the rectangle inside which the researcher can describe the topic, and from which the Idea Coach returns with 15 questions. The actual text of the request is: *I am a doctor treating obesity in children. How do I talk to the children to make them understand?*

Table 1: Questions emerging from the Idea Coach which address the topic. The topic is provided by the researcher.

<p>Topic: I am a doctor treating obesity in children. How do i talk to the children to make them understand</p> <p>Topic Questions:</p> <p>Do you know why it's important to maintain a healthy weight?</p> <p>How do you think being overweight can affect your overall health?</p> <p>What types of activities do you enjoy doing to stay active?</p> <p>Can you identify some unhealthy eating habits that may contribute to weight gain?</p> <p>How can we make healthier food choices when eating out at restaurants?</p> <p>What are some enjoyable ways to include physical activity into your daily routine?</p> <p>Do you think it's important to limit screen time? Why or why not?</p> <p>How can we involve your family or friends in making healthier choices?</p> <p>Can you think of some foods that are high in nutrition and low in calories?</p> <p>Do you know any strategies to control portion sizes during meals and snacks?</p> <p>Have you ever felt judged or teased because of your weight? How did that make you feel?</p> <p>How can we overcome challenges or obstacles that prevent us from making healthier choices?</p> <p>Can you think of ways we can make healthy snacks more appealing and enjoyable for you?</p> <p>What are your thoughts on keeping a food and activity journal to track progress towards healthier habits?</p> <p>How can we make sure that making healthier choices doesn't feel like a punishment or a chore?</p>

After the researcher receives the various sets of 15 questions, the next task is for the researcher to provide four questions. These questions can be taken from those suggested by the AI, either 'as is' or edited to tailor the 'language' and 'style' of the answer. The researcher may also contribute questions. Quite often the questions shown in Table 1 have to be modified, not so much for the respondent who never sees the questions, but rather for the AI to provide the proper format of the answer.

Table 2 shows the modified question, edited by the researcher, and submitted to the Idea Coach. In turn, for each question in Table 2, Idea Coach returned with sets of 15 answers, formatted in the way request by the researcher. Table 2 also shows the four answers to each question, these answers having been returned by AI (Mind Genomics' Idea Coach), and then slightly edited to make them flow more easily. The rationale was to generate answers that could be given to the Idea Coach programmed to act as a human respondent.

Running the Study and Preparing the Data for Regression Modeling and for Clustering

Following the creation of the edited answers, the Mind Genomics platform was instructed to run a study with 301 synthetic respondents. The program was instructed to create a panel comprising approximately equal numbers of the three types of synthetic respondents defined by their job (receptionist, doctor with 10-years-experience, nurse practitioner). Then the Mind Genomics program presented 24 vignettes to the synthetic respondent, defining WHO the respondent is, defining the SCALE (Table 3), and then one vignette at a time to the AI 'respondent' with the request to rate the vignette on the two-sided scale by choosing one rating point. This means that the AI had to consider the vignette from the complex of how the patient would feel and how the doctor would feel. The Mind Genomics platform recorded the information about the respondent, the vignette, the rating, and the time elapsed for the synthesized respondent to rate the vignette.

Table 2: The four questions and the four answers to each question. The questions and the answers were edited slightly to make them understandable to human beings, but the language and meaning was maintained.

	Question A: As a doctor, approach children with empathy, understanding, and a non-judgmental attitude about their weight. Answer this question in 15 words, in a way that I can show it on a board in the classroom to a child age 9-11
A1	Weight is different for everyone and doesn't define your worth or value as a person.
A2	We can find fun ways to stay active and make healthy choices together.
A3	Let's talk about ways to keep your body strong and full of energy
A4	Remember, everyone grows and develops at their own pace.
	Question B: Talk to the child about some of the problems that we recognize. Tell the child that here are some real problems that make it difficult to be active. Answer this question in 15 words, in a way that I can show it on a board in the classroom to a child age 9-11
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.
B4	Here is a real problem we hear about: Inadequate knowledge about the importance of exercise and its benefits.
	Question C: .How can I support you in setting realistic goals for weight loss and creating a healthier lifestyle? Answer this question in 15 words, in a way that I can show it on a board in the classroom to a child age 9-11
C1	Here is part of our program designed for you: Share inspiring success stories of people who achieved their weight loss goals.
C2	Here is part of our program designed for you Motivate with positive and encouraging words.
C3	Here is part of our program designed for you Offer support in overcoming any obstacles along the way.
C4	Here is part of our program designed for you Involve the child in grocery shopping and meal planning.
	Question D: Can you tell me about your feelings and emotions related to your weight? Answer this question in 15 words, in a way that I can show it on a board in the classroom to a child age 9-11
D1	Here are some of the feelings we will help you deal with in our program: "Sometimes I may feel self-conscious if others tease or make comments about my weight."
D2	Here are some of the feelings we will help you deal with in our program: "I want to focus on being healthy rather than just how much I weigh."
D3	Here are some of the feelings we will help you deal with in our program: "I know that my worth isn't defined by the number on the scale."
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."

Table 3: The preliminary question for assignment into respondent job (top) and then the 'two-sided rating scale (bottom).

Preliminary question: Who are you Possible answers: 1=I am a receptionist in the office, talking to someone coming into the doctor, but just in the process of talking to the patient. 2=I am a doctor with 10 years of experience 3=I am a nurse practitioner with 10 years of experience
Rating question: What is the likely outcome of this visit Scale Points 1=Patient not motivated and medical person feels the wrong things were said 2=Patient not motivated but medical person feels the right things were said 3=Can't make a decision 4=Patient motivated but medical person feels the wrong things were said 5=Patient motivated and medical person feels the right things were said

Once the data was collected for a ‘respondent’ the AI returned the raw data to the Mind Genomics platform to create the database shown as an Excel file in Figure 2. Each respondent generates 24 rows of data, one row for each of the 24 vignettes. Figure 2 is divided into several sections, representing different aspects of the data and of the analysis. As we follow the structure of the file, we must keep in mind that as the data is being transferred to the database, there are preparatory transformations occurring in ‘real time,’ these transformation necessary for the analysis.

1. Row-for the entire database. There are 301 respondents, each with 24 rows of data, generating 7224 rows, each row corresponding to a specific respondent and a specific vignette.
2. Study and panelist number.
3. The AI assignment of the respondent to one of three groups (receptionist, doctor with 10-years-experience, nurse practitioner) by the membership of the synthesized respondent into one of two mind-sets, and then into one of three mind-sets.
4. Columns showing test order and ‘half’ defined as 1 for test order 1-12, and 2 for test order 13-24, respectively.
5. The 16 elements, A1-D4, coded as ‘1’ when present in the vignette, coded as ‘0’ when absent from the vignette.
6. Rating, new variables R54x (patient motivated), R12x (patient not motivated), R52x (medical person feels the right thing was said), R41x (medical person thinks the wrong things were said), R3x (cannot answer), and RT (response time to

the nearest 10th of a second. The new variables were created by adding the appropriate variables. In no case does the creation of a new variable produce any number other than 0 or 1. After these variables were created a vanishingly small random number was added to each new variable (viz., R54x, R12x, R52x, R41x, R3) to ensure that these newly created variables possessed minimal variability as required for analysis by regression.

Using Regression Modeling to Link the 16 Elements to the Newly Created Binary Variables

The first analysis looks at the average ratings of the newly created binary variables (e.g., R54x, Patient Motivated). The question is whether the instruction to the AI to assume a certain persona (e.g., receptionist) has an effect on the average ratings across all the vignettes in which the AI assumed that it was a ‘receptionist’. Table 4 shows that the average ratings for a specific binary variable are quite similar across the three different ‘AI personas’. Our first conclusion leaves us with the concern that AI may not be able to perform the task in a way which makes sense.

The final analysis for this first run used OLS (ordinary least squares regression, to relate the 16 elements to the binary transformer rating. The equation is: Transformed Binary Variable= $k_1(A1) + k_2(A2) + \dots + k_{16}(D4)$. That is, the regression model uses the data to show how the presence or absence of the element drives the binary variable. The least-squares regression can be run for the total panel, and for all respondents defined by the AI synthesizer as being ‘receptionist’, ‘doctor with 10-years-experience’, or nurse practitioner.

Row	Study	Panelist	Quest1	TWO MS	THREE MS	TestOrder	Half	A1	A2	A3	A4	B1	B2	B3	B4	C1	C2	C3	C4	D1	D2	D3	D4	Rating	R54x	R12x	R52x	R41x	R3x	RT
1	11192023.Hov	1	2	2	2	1	1	0	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0	3	0	1	0	0	100	4
2	11192023.Hov	1	2	2	2	2	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	3	1	1	0	0	100	3.3
3	11192023.Hov	1	2	2	2	3	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	3	0	0	0	0	100	1
4	11192023.Hov	1	2	2	2	4	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0	0	2	1	100	100	0	0	2.3
5	11192023.Hov	1	2	2	2	5	1	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	3	1	1	0	0	100	0.7
6	11192023.Hov	1	2	2	2	6	1	0	1	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1	1	100	0	100	0	4
7	11192023.Hov	1	2	2	2	7	1	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	1	3	1	0	0	0	100	1.6
8	11192023.Hov	1	2	2	2	8	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	2	0	101	100	0	0	0.7
9	11192023.Hov	1	2	2	2	9	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0	100	2.1
10	11192023.Hov	1	2	2	2	10	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	0	3	1	0	0	0	100	3
11	11192023.Hov	1	2	2	2	11	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	2	1	101	100	0	0	1.2
12	11192023.Hov	1	2	2	2	12	1	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	3	1	1	0	0	100	3.2	
13	11192023.Hov	1	2	2	2	13	2	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0	3	1	1	0	0	100	3.7
14	11192023.Hov	1	2	2	2	14	2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	1	2	0	101	100	0	0	1.3

Figure 2: Example of database showing the actual information generated by the AI, and the set of responses to each vignette.

Table 4: Average rating for five binary variables and ‘response time’ for three respondent personas created by AI and set to the task of rating the vignette. The data come from the averages of the newly created binary variables. Each average comes from approximately 2400 numbers, assuming that each synthesized persona comprised 100 respondents, each respondent evaluating 24 vignettes.

	Patient Motivated R54x	Patient Not Motivated R21x	Right Thing Said R52x	Wrong Thing Said R41x	Can’t Answer R3x	Response Time RT
Receptionist	32	24	26	30	44	2.3
Doctor	30	26	28	28	45	2.2
Nurse Practitioner	32	25	27	29	44	2.2

Table 5 shows the regression coefficients, computed as if the ratings and therefore the binary variables (e.g., R54x) came from people. Relevant coefficients are shown in shaded cells. There may be a number of stories within the data, but no simple organizing principle.

Experience with ratings generated by people produce the same story, namely some strong performers but not many, and stories that could be told, but seem too isolated.

Table 5: Regression coefficients for the key groups (receptionist, doctor with 10-years-experience, and nurse practitioner. The regression coefficients were estimated by OLS regression based upon the ratings of the vignette assigned by artificial intelligence.

		Patient Motivated R54x			Right thing said R52x		
		Receptionist	Doctor	Nurse Practitioner	Receptionist	Doctor	Nurse Practitioner
A1	Weight is different for everyone and doesn't define your worth or value as a person.	11	12	10	6	5	8
A2	We can find fun ways to stay active and make healthy choices together.	14	7	9	10	5	9
A3	Let's talk about ways to keep your body strong and full of energy	11	12	5	9	4	5
A4	Remember, everyone grows and develops at their own pace.	13	10	9	9	0	5
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.	12	6	9	9	11	11
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	11	6	10	12	9	8
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.	11	5	6	9	11	11
B4	Here is a real problem we hear about: Inadequate knowledge about the importance of exercise and its benefits.	7	5	8	7	9	7
C1	Here is part of our program designed for you: Share inspiring success stories of people who achieved their weight loss goals.	9	11	11	6	9	9
C2	Here is part of our program designed for you Motivate with positive and encouraging words.	7	13	11	9	8	8
C3	Here is part of our program designed for you Offer support in overcoming any obstacles along the way.	10	11	12	8	7	8
C4	Here is part of our program designed for you Involve the child in grocery shopping and meal planning.	9	12	10	7	12	8
D1	Here are some of the feelings we will help you deal with in our program: "Sometimes I may feel self-conscious if others tease or make comments about my weight."	9	5	12	5	9	7
D2	Here are some of the feelings we will help you deal with in our program: "I want to focus on being healthy rather than just how much I weigh."	6	10	8	7	10	6
D3	Here are some of the feelings we will help you deal with in our program: "I know that my worth isn't defined by the number on the scale."	5	7	9	7	10	6
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."	6	11	12	9	10	3
		Patient Not Motivated R12x			Wrong Thing Said R41x		
		Receptionist	Doctor	Nurse Practitioner	Receptionist	Doctor	Nurse Practitioner
A1	Weight is different for everyone and doesn't define your worth or value as a person.	7	4	8		12	8
A2	We can find fun ways to stay active and make healthy choices together.	9	9	7	12	10	7
A3	Let's talk about ways to keep your body strong and full of energy	11	5	7	12	13	6
A4	Remember, everyone grows and develops at their own pace.	10	2	4	14	11	7
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.	7	12	10	10	7	7
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	12	9	8	10	6	9
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.	7	11	12	8	4	7
B4	Here is a real problem we hear about: Inadequate knowledge about the importance of exercise and its benefits.	8	11	8	8	7	9
C1	Here is part of our program designed for you: Share inspiring success stories of people who achieved their weight loss goals.	4	8	6	6	9	8
C2	Here is part of our program designed for you Motivate with positive and encouraging words.	6	8	5	4	13	8
C3	Here is part of our program designed for you Offer support in overcoming any obstacles along the way.	5	7	6	7	10	9
C4	Here is part of our program designed for you Involve the child in grocery shopping and meal planning.	8	11	7	10	11	9

D1	Here are some of the feelings we will help you deal with in our program: "Sometimes I may feel self-conscious if others tease or make comments about my weight."	4	7	7		7	2	11
D2	Here are some of the feelings we will help you deal with in our program: "I want to focus on being healthy rather than just how much I weigh."	7	4	8		6	4	10
D3	Here are some of the feelings we will help you deal with in our program: "I know that my worth isn't defined by the number on the scale."	8	6	7		7	3	10
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."	6	7	6		3	8	14
		Can't Make Decision R3x			Response Time RT			
		Receptionist	Doctor	Nurse Practitioner		Receptionist	Doctor	Nurse Practitioner
A1	Weight is different for everyone and doesn't define your worth or value as a person.	11	12	13		0.6	0.7	0.8
A2	We can find fun ways to stay active and make healthy choices together.	6	11	12		0.7	0.6	0.7
A3	Let's talk about ways to keep your body strong and full of energy	7	10	15		0.6	0.5	0.7
A4	Remember, everyone grows and develops at their own pace.	7	17	18		0.7	0.6	0.8
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.	13	13	14		0.6	0.7	0.8
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	8	15	12		0.6	0.7	0.6
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.	12	15	12		0.6	0.8	0.7
B4	Here is a real problem we hear about: Inadequate knowledge about the importance of exercise and its benefits.	16	14	17		0.7	0.7	0.7
C1	Here is part of our program designed for you: Share inspiring success stories of people who achieved their weight loss goals.	18	12	13		0.7	0.7	0.7
C2	Here is part of our program designed for you Motivate with positive and encouraging words.	18	9	12		0.7	0.7	0.5
C3	Here is part of our program designed for you Offer support in overcoming any obstacles along the way.	16	14	12		0.7	0.7	0.6
C4	Here is part of our program designed for you Involve the child in grocery shopping and meal planning.	13	9	13		0.7	0.7	0.7
D1	Here are some of the feelings we will help you deal with in our program: "Sometimes I may feel self-conscious if others tease or make comments about my weight."	15	17	10		0.6	0.5	0.6
D2	Here are some of the feelings we will help you deal with in our program: "I want to focus on being healthy rather than just how much I weigh."	15	13	11		0.6	0.6	0.5
D3	Here are some of the feelings we will help you deal with in our program: "I know that my worth isn't defined by the number on the scale."	13	16	11		0.7	0.6	0.6
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."	16	11	11		0.6	0.6	0.5

A better way to look for ‘humanness’ in the data may be to search for strong-performing elements for each. As a further step, one might present these strong performing elements to others, and ask them how they feel about the motivating power of the elements, or how they feel about the elements in terms of what might be expected from someone in the medical profession. As a first approximation, there are no strong surprises in Table 6, which shows the strongest elements for each newly created binary variable (e.g., R54x, Patient Motivated), when the AI took on three of the personas.

The final set of analyses, clustering, divides the respondents into mutually exclusive and exhaustive groups, called ‘clusters’ by statisticians, ‘segments’ by consumer researchers, and ‘mind-sets’ in the language of Mind Genomics. The underlying notion, foundational for Mind Genomics, is that ‘people’ differ systematically in the way they make decisions about the world of the everyday. The plethora of different products, services, even layouts of towns and the styles of houses and their decorations clearly announce these differences in judgment. Rather than considering this person-to-person variability to be an unwanted secondary factor, noise, in an otherwise simple

world, Mind Genomics looks for organizing principles, so-called mind-sets. These mind-sets are derived empirically from a study of the differences among people in how those people evaluate the world of the everyday, at the granular level, not the 20,000-foot level. A variety of papers have been published on the use of Mind Genomics to identify these mind-sets in various situations and for various products.

The question for this study is whether mind-sets can be uncovered when we deal with synthetic respondents. The process to discover these mind-sets will certainly work with the data generated by AI because the input data needed to create the mind-sets is simply the set of coefficients, one per respondent, as shown in Table 5, specifically in this study for coefficients where R54x (patient motivated) is the dependent variable.

The process for creating mind-sets follows strict mathematical rules. The clustering is totally insensitive to the ‘meaning’ of the elements and does not care from where the coefficients came. Mind-sets emerge after the researcher develops the equation for each of synthesized respondent, puts the data into a matrix of 301 rows (one row for each synthesized respondent, 16 columns (one column for

Table 6: Strongest performing elements for three AI personas (receptionist, doctor, nurse practitioner) on four newly created binary variables.

Patient Motivated R54x		
	1. Receptionist	
A2	We can find fun ways to stay active and make healthy choices together.	14
2. Doctor		
C2	Here is part of our program designed for you Motivate with positive and encouraging words.	13
3. Nurse Practitioner		
C3	Here is part of our program designed for you Offer support in overcoming any obstacles along the way.	12
..	Here are some of the feelings we will help you deal with in our program: "Sometimes I may feel self-conscious if others tease or make comments about my weight."	12
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."	12
Patient NOT Motivated R12x		
	1. Receptionist	
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	12
2. Doctor		
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.	12
3. Nurse Practitioner		
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.	12
Right Thing Said R52x		
	1. Receptionist	
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	12
2. Doctor		
C4	Here is part of our program designed for you Involve the child in grocery shopping and meal planning.	12.
3. Nurse Practitioner		
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.	11.
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.	11.
Wrong Thing Said R41x		
	1. Receptionist	
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	12
2. Doctor		
C2	Here is part of our program designed for you Motivate with positive and encouraging words.	13
A3	Let's talk about ways to keep your body strong and full of energy	13
3. Nurse Practitioner		
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."	14

each of the 16 coefficients from regression). The k-means clustering program [9] then creates groups of synthesized respondents whose patterns are similar across the 16 coefficients.

The process of creating the mind-sets is straightforward, using clustering. The process follows these steps:

1. Create the basic equation for each respondent: $R54x = k_1A1 + k_2A2 \dots k_{16}D4$. The equation has 16 coefficients, estimated from the 24 cases generated for each synthetic respondent.
2. Estimate the ‘distance’ or ‘dissimilarity’ between each pair of the 301 respondents by computing the Pearson Correlation

(R). Then create the new distance parameter, (1-R). The value (1-R) is 0 when the two sets of coefficients are parallel, meaning that the two items, our synthetic respondents, show identical patterns. The value (1-R) is 2 when the two sets of coefficients are opposite.

3. Use clustering to assign respondents to one of two clusters, or later one of three clusters. The criterion is that the variability within a cluster should be ‘small’ because the respondents show similar patterns of coefficients. In contrast, the variability across the centroids of the clusters should be large because the clusters are different groups.

- Once the respondents are assigned to the appropriate cluster (which of the two clusters, which of the three clusters), recompute the equation using the data from all respondents in a cluster.

Table 7 shows the performance of the elements for the Total Panel, for the two mind-sets, and for the three mind-sets. Just looking at the Table shows a number of shaded cells with coefficients of 21 or higher. High coefficients by themselves do not suffice, however. Rather, the coefficients must ‘tell a story’, and allow for interpretation.

The first clustering creating two mind-sets produces easily interpreted mind-sets.

Mind-Set 1 of 2-Focus on Activity and Vitality

We can find fun ways to stay active and make healthy choices together.

Let’s talk about ways to keep your body strong and full of energy

Mind-Set 2 of 2-Focus on What Specifics to Do

Here is part of our program designed for you Motivate with positive and encouraging words.

Here is part of our program designed for you Involve the child in grocery shopping and meal planning.

Here is part of our program designed for you: Share inspiring success stories of people who achieved their weight loss goals.

In contrast, the three mind-set-solution produces elements with higher coefficients, but the underlying pattern is hard to interpret.

Our final analysis considers the performance of the elements using the IDT, Index of Divergent Thought [10]. The IDT attempts to quantify the degree to which the clustering generates truly different groups of people based upon the coefficients. Table 8 shows the computations. Optimal levels of the IDT are found in the range of 70-75. Higher values of the IDT (e.g., 80 or higher) mean that there are many high coefficients but not dramatically patterns of coefficients across the mind-sets Lower values of the IDT (e.g., 60 or lower) mean that there are many low coefficients, and once again the pattern of differences in coefficients across mind-sets is simply not dramatic.

For this study, and for others not reported here, the pattern of low values for the IDT continues to emerge when we deal with synthetic respondents. Here it is 47. Simply stated, the synthetic respondents may ‘work’ but do not yet produce dramatically stronger performance. As a starting point, however, it is gratifying to see that initial exploration into the AI does produce interpretable mind-sets.

Discussion and Conclusions

The emerging interest in AI generated respondents, so-called synthetic respondents, provides a new area of opportunity for the equally emerging science of Mind Genomics. As shown here, it is straightforward to craft a series of prompts to AI for a specific topic, these prompts being a description of ‘WHO’ the person is (our three

Table 7: Coefficients emerging from clustering the 301 synthetic respondents into two mind-sets and then three mind-sets, based on the pattern of the coefficients for R54x (Motivates).

		Total	Mindset 1 of 2	Mindset 2 of 2	Mindset 1 of 3	Mindset 2 of 3	Mindset 3 of 3
	Base (number of respondents in this group)	301	133	168	76	129	96
	Mind-Set 1 of 2-Focus on activity and vitality						
A2	We can find fun ways to stay active and make healthy choices together.	10	21	1		5	28
A3	Let’s talk about ways to keep your body strong and full of energy	9	21		1	2	24
	Mind-Set 2 of 2-Focus on what specifics to do						
C2	Here is part of our program designed for you Motivate with positive and encouraging words.	11		22		27	
C4	Here is part of our program designed for you Involve the child in grocery shopping and meal planning.	10		21		25	
C1	Here is part of our program designed for you: Share inspiring success stories of people who achieved their weight loss goals.	11		21		24	1
	Not strong for either of the two mind-sets						
C3	Here is part of our program designed for you Offer support in overcoming any obstacles along the way.	11		20	4	23	1
B1	Here is a real problem we hear about: Lack of safe outdoor spaces.	9	7	10	24	7	
B2	Here is a real problem we hear about: Busy schedules and lack of free time to play.	9	7	10	20	9	
B3	Here is a real problem we hear about: Lack of parental support or encouragement for being active.	7	6	8	19	6	
B4	Here is a real problem we hear about: Inadequate knowledge about the importance of exercise and its benefits.	7	4	8	17	5	
A1	Weight is different for everyone and doesn’t define your worth or value as a person.	11	19	5	1	7	25
D2	Here are some of the feelings we will help you deal with in our program: "I want to focus on being healthy rather than just how much I weigh."	8	12	5	19	3	8
A4	Remember, everyone grows and develops at their own pace.	10	18	4		7	25
D4	Here are some of the feelings we will help you deal with in our program: "It's okay to have insecurities, but it's more important to love and accept ourselves."	10	16	4	21	1	13
D1	Here are some of the feelings we will help you deal with in our program: "Sometimes I may feel self-conscious if others tease or make comments about my weight."	9	15	4	18		13
D3	Here are some of the feelings we will help you deal with in our program: "I know that my worth isn't defined by the number on the scale."	7	15	1	17		13

Table 8: Calculation of the IDT, Index of Divergent Thought, for the synthesized data. The IDT is based on clustering coefficients estimated for R54x, patient is motivated.

	Total	Mind-Set 1 of 2	Mind-Set 2 of 2	Mind-Set 1 of 3	Mind-Set 2 of 3	Mind-Set 3 of 3
Base (number of respondents in this group)	301	133	168	76	129	96
Sum Total of Squares of All Positive Coefficients	1419	2567	2194	3059	2787	3183
Average of All Positive Coefficients (sum coefficients/base)	5	19	13	40	22	33
Number of respondents in the column (in the subgroup)	301	133	168	76	129	96
Weight=Proportion of the respondents in the subgroup	0.33	0.15	0.19	0.08	0.14	0.11
Weighted Total (Weight x Sum Total of Squares of All Positive Coefficients)	468	385	417	245	390	350
Final Score=sum of weighted total	2255					
IDT=Square root of final score	47					

medical professionals), how the person is to ‘JUDGE’ (the rating scale), and finally ‘WHAT’ the person is to judge (the vignette).

The Mind Genomics task is difficult for people. It presents combinations of elements which may or may not fit together, but which at least do not contradict each other. Yet, human beings can do it. What is remarkable is that AI can do the task, perhaps not as well as people because of the lower coefficients, but nonetheless AI can do the job.

What is most remarkable is that AI with synthetic respondents can deal with the two-sided scale, one side for motivating or not motivating the patient, the second side for the right versus the wrong thing said.

‘Looking to the meaning of the data, we can focus on both the philosophical issue of the ‘Turing Test’ and the use of the approach to create a body of knowledge for teaching. The issue of the Turing Test is a well known one in philosophy. Quite simply, can we create a test which to figure out whether a machine is a machine or a human being. The data here suggest that the Mind Genomics process can be reasonably mimicked by a machine, with the answers ‘making sense.’

Following quickly on the heels of the foregoing question is whether or not the Mind Genomics system can be engineering to become a teaching/learning system, wherein we design the persona to have a variety of emotional and other issues, and then evaluate the response of alternative descriptions of treatments. The sheer number of papers dealing with this issue of learning about interactions between groups of people is heartening [11-13]. This second avenue is likely to be the one more interesting, and ultimately more fruitful to researchers, to philosophers and to society alike.

Acknowledgment

The senior author, HRM, dedicates this paper to the memory of a mentor whom he never met, but who was instrumental in his thinking since 1965. The mentor is the late Harvard University professor, Anthony Gervin Oettinger of blessed memory, who laid the foundation of HRM’s interest in artificial intelligence. It was Professor Oettinger who planted the seeds of this paper almost six decades ago, in his offer to have the author participate in the 1960’s Harvard project, TACT, Technical Aids to Creative Thought. Thank you Tony. This paper is for you.

References

- Dave T, Athaluri SA, Singh S (2023) ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in Artificial Intelligence* 6: p.1169595. [crossref]
- Campbell RT, Hudson CM (1985) Synthetic cohorts from panel surveys: An approach to studying rare events. *Research on Aging* 7: 81-93. [crossref]
- Dang HAH, Dabalén AL (2019) Is poverty in Africa mostly chronic or transient? Evidence from synthetic panel data. *The Journal of Development Studies* 55: 1527-1547.
- Dang HAH, Lanjouw PF (2018) Poverty dynamics in India between 2004 and 2012: Insights from longitudinal analysis using synthetic panel data. *Economic Development and Cultural Change* 67: 131-170.
- Gofman A, Moskowitz H (2010) Isomorphic permuted experimental designs and their application in conjoint analysis. *Journal of Sensory Studies* 25: 127-145.
- Moskowitz HR (2012) ‘Mind genomics’: The experimental, inductive science of the ordinary, and its application to aspects of food and feeding. *Physiology & Behavior* 107: 606-613.
- Moskowitz HR, Gofman A (2007) *Selling Blue Elephants: How to Make New Products That People Want Before They Even Know They Want Them*. Pearson Education.
- Moskowitz HR, Martin D (1993) *How Computer Aided Design and Presentation of Concepts Speeds up the Product Development Process*. Proceedings of the ESOMAR Congress Copenhagen.
- Likas A, Vlassis N, Verbeek JJ (2003) The global k-means clustering algorithm. *Pattern Recognition* 36: 451-461. [crossref]
- Todri A, Papajorgij P, Moskowitz H, Scalera F (2020) Perceptions regarding distance learning in higher education, smoothing the transition. *Contemporary Educational Technology* 13: p.ep287.
- Berry DC, Michas IC, Gillie T, Forster M (1997) What do patients want to know about their medicines, and what do doctors want to tell them?: A comparative study. *Psychology and Health* 12: 467-480.
- Collins J, Farrall E, Turnbull DA, Hetzel DJ, Holtmann G, et al. (2009) Do we know what patients want? The doctor-patient communication gap in functional gastrointestinal disorders. *Clinical Gastroenterology and Hepatology* 7: 1252-1254. [crossref]
- Kay M, Mitchell G, Clavarino A (2010) What do doctors want? A consultation method when the patient is a doctor. *Australian Journal of Primary Health* 16: 52-59. [crossref]

Citation:

Moskowitz HR, Rappaport S, Morford T, Mulvey M, Tomaro K, et al. (2023) Advancing Mind Genomics: Using AI (Artificial Intelligence) to Generate Topics, Questions, Messages as Answers, and ‘Synthesized’ Minds (Respondents). *Psychol J Res Open* Volume 5(8): 1-9.