

Review Article

The Quality of Ideas When AI (Artificial Intelligence) is Used as a Coaching Device

Howard Moskowitz^{1*}, Stephen Rappaport² and Yehoshua Deitel³

¹Mind Genomics Associates, Inc. White Plains, NY, USA

²SDR Consulting, Inc., Norwalk, CT, USA

³Sifra Digital, Jerusalem, Israel

*Corresponding author: Howard Moskowitz, Mind Genomics Associates, Inc. White Plains, NY, USA

Received: June 05, 2023; Accepted: June 12, 2023; Published: June 19, 2023

Abstract

Four studies explored the subjective responses to text generated by AI versus text generated by a single individual. Each study focused on a topic (local pizza shop; local cleaner; local dentist; local school board). Each study comprised four sets of four elements each, viz., 16 elements, the elements being answers to simple questions pertaining to the topic. Each study comprised 30-31 respondents. The evaluation showed no decided advantage for AI generated information over human generated information when the rating was assigned on a simple 5-point Likert scale. The test process provides a way of inserting human evaluation into the text generated by people or by AI, and in turn opens up the potential of a standardized way to evaluate AI text using people.

Introduction

Today's focus on hot new technologies must, by sheer popularity, include AI, artificial intelligence. AI is high on the Gartner hype cycle (Dedehayir & Steinert, 2016), perhaps due to its romanticized connotations as much to its performance. This paper is simply a study of how well AI can help a person think, or solve a problem, when AI is used as a 'coaching device.' At the same time this paper is being written in a time of ferment. Two decades ago Hill et. al., (2005), reported that "people communicated with the chatbot for longer durations (but with shorter messages) than they did with another human. Additionally, human-chatbot communication lacked much of the richness of vocabulary found in conversations among people and exhibited greater profanity. These results suggest that while human language skills transfer easily to human-chatbot communication, there are notable differences in the content and quality of such conversation". The big story of 2022 is that AI has developed so quickly and powerfully that it is being used (or really misused) by students to write their college essays (Fyfe, 2022.).

The origins of this study come from three decades of work on systematic exploration of people's responses to ideas, when these ideas are presented in combinations, in vignettes, with these vignettes created by experimental design. The origin of the work is the emerging science of Mind Genomics (Moskowitz et. al., 2012). Mind Genomics focuses on how people make decisions in the world of ordinary experience. The underlying approach of Mind Genomics is to choose a topic, then ask a set of four questions which present the topic in a story sequence, choose four answers to each question, creating 16 elements (element=answer). Finally, the Mind Genomics program (www.bimileap.com) creates combinations of these elements according to an underlying set of recipes, a so-called experimental design (Gofman

& Moskowitz, 2010), presents these combinations to the respondent, obtains ratings, and uses regression analysis to deconstruct the responses into the contributions of the different elements.

The foregoing approach has been templated and is now easy to use. One recurring observation over the decades of use is that people simply freeze when they are instructed to create a set of four questions. That is, people may know how to answer questions, but again and again it is becoming clear that people don't know how to frame questions. In other words, it may be that people simply don't know how to think in a critical way, at least when critical thinking is defined in part as the ability to ask a set of coherent questions.

Introducing AI into Mind Genomics

Beginning around 2010, Mind Genomics was put on a path of developing a true DIY (do it yourself) system, with automatic set-up, automated recruiting using a panel company as a partner, and automated reporting. The effort finally materialized in 2017, with the set-up screens. The objective of the DIY was to have the researcher identify a topic for the study, then provide four questions, and finally four answers for each question.

Years of experience gave the authors the preparation to provide questions and answers with a fair degree of ease. It seemed to the authors that almost anyone should be able to think of four questions about a topic, but the authors were slowly disabused of that notion when users had to be 'coached', brought along, and only after the coaching and talking were the users comfortable. The unhappy reality was that the reluctance, or perhaps the real inability, to come up with questions was hindering the adoption of the Mind Genomics platform. Users simply abandoned the Mind Genomics effort, often

after choosing a topic, and being confronted with what had seemed to be a perfect simple task of choosing questions.

During the many years of developing Mind Genomics, author HRM was asked many times whether the questions and answers could be put in somehow 'automatically.' The history of those questions would make an interesting article in itself, but the main thing to emerge was that people simply did not know how to think about topics in a critical way. Answering questions was very easy for people. In contrast, framing questions about a topic was off-putting. It seemed at first that people were being lazy, but the reality began to emerge that people were simply recognizing the need for a coach, for something to help them, preferably not a person.

The opportunity to create a 'coaching' system emerged during the middle of the past year, in 2022. At that time the Mind Genomics approach had become virtually a DIY system, as long as the respondent knew how to create the right questions, create the answers, create the orientation page, and the rating scale. The orientation page and rating scales were easy to learn and master. Creating questions and answers remained elusive, depending as much on the person's confidence and topic-specific knowledge as it did on the person's attitude towards these open ended explorations.

It would be the available of AI as an interface which provided the opportunity for enhancing Mind Genomics, perhaps removing the barrier of having to think of questions for a given topic. The underlying idea was to provide a 'coaching mechanism', effectively a device in which researchers could elaborate on a topic, with AI returning with up to 30 questions. The elaboration could be done several times, either with the same text or with different variations, until the researcher came up with questions which seemed relevant. Some or all of these questions could be selected, and 'dropped' into the question boxes. The researcher was free to edit the questions, and to provide questions that she or he wished to add, as long as the number of questions came to four. The second part of the coach allowed the researcher to select the coach, receive up to 15 answers to the question, and again repeat the process, selecting the best answers. The researcher could edit the answers, as well as incorporate her or his own answers.

It was the introduction of AI in this way, viz., as a coach, which reduced some of the hesitancy, making Mind Genomics more of a cut and paste operation in which the thinking was encouraged, but not made into a stumbling block. The researcher could be a child, even one eight years old or so. The fear factor of thinking was replaced by the excitement of the new information to appear, new questions to check, new answers to discover and select. The task became fun, at least to children. And, for adults, the delight may not have been so obvious, but the hesitation appeared to have been reduced, in some occasions dramatically.

AI Generated Ideas Versus Self-generated Ideas – Do They Differ in Quality?

The genesis of this paper emerged from ongoing conversations about the use of AI in studies. The question continue to emerge regarding the 'quality' of the ideas emerging from AI. The consumer researcher business continues to feature advertisements about AI in

consumer research, although the underlying methods, viz. 'under the hood' are never revealed, nor are metrics. Yet despite the opacity of AI, the pattern is clear. It may only be a matter of time before AI catches up with the human being, and actually surpasses people (Grace et al. 2018). They write in stark terms that "AI will outperform humans in many activities in the next ten years, such as translating languages (by 2024), writing high-school essays (by 2026), driving a truck (by 2027), working in retail (by 2031), writing a bestselling book (by 2049), and working as a surgeon (by 2053). Researchers believe there is a 50% chance of AI outperforming humans in all tasks in 45 years and of automating all human jobs in 120 years "

A consequence of the popularity of AI, at least in the discussions among market researchers has led to the question of how one might measure the quality of ideas emerging from AI. Would the ideas be really better, producing insights that were the equal of good ideas produced by people, and hopefully better ideas? Or, when subject to standard researchers, such as Mind Genomics, would AI produced ideas perform worse than or equal to, but definitely not better than ideas of smart people. It was this last idea, head to head comparison to the performance of ideas generated by AI versus by people which gave rise to this paper.

Mind Genomics and the Head to Head Comparison of Ideas

Mind Genomics allows the researcher to measure the quality of the ideas, without anyone except the researcher knowing the source. The research project emerged as simple to do. In a few words, the idea was to choose a simple topic, and run the study exactly the same two times. The first time the researcher would create the study with no help at all. The second time the research would select the elements, the raw material, from Idea Coach. The same number of respondents would participate in the parallel studies, and the result compared in terms of performance.

There is no direct, absolute, non-subjective way to measure the quality of ideas. Thus, the ideas which emerge from one's unaided efforts cannot be 'objectively' compared to the ideas which emerge from AI. AI may produce many more ideas, but on the basis of what can we assign values to these ideas? There is the old adage, a bit judgmental, and somewhat contemptuous, namely 'GIGO, garbage in, garbage out.' Given the prevalence of GIGO thinking, is possible to create a system to measure the value of ideas.

Mind Genomics may provide a solution to the above-mentioned problem. Keep in mind that there are two parts of the Mind Genomics exercise. The first part comes from the selection of the elements, whether from the person or from the combination of AI as the provider of the elements and the person as the selector of the elements. The second part comes from the evaluation of the elements, albeit combinations of elements, not single elements alone. The evaluation is one by real people who do not know how the elements were generated, rating each combination on a simple Likert scale, viz. an anchored scale. With this separation of idea generation from idea evaluation, it may be possible to measure the human subjective response to the ideas, and by so doing compare the performance of ideas generated by people to ideas generated by AI.

The Mind Genomics Process

Mind Genomics is an emerging science, focusing on the way people make decisions (REF). Mind Genomics grew out of a combination of three different disciplines; statistics, psychophysics, and consumer research, respectively. The underlying process has been documented a number of times (REF). In summary, the process follows these steps:

Step 1 – Define the topic, generate four questions, and for each question generate four answers. The term ‘element’ will be used instead of the term ‘answer.’ In effect, there are 16 elements in a Mind

Genomics study. These questions and answers may be generated by the researcher, or may be generated using artificial intelligence. In this study, four of eight studies used questions and questions generated by the researcher, and the other four studies, with matching topics, used questions and answers generated by OpenAI’s language model “text-davinci-003” (<https://beta.openai.com/docs/models/gpt-3>).

The eight studies summarized in Table 1 were run with 30-31 respondents, all in the state of Connecticut, USA, with respondents aged 18-49 for the pizza studies, and 25-54 for the dry cleaner, dentist and school board candidate studies, respectively.

Table 1: Positive coefficients for the eight studies, showing the results for the Total panel, and for the three mind-sets extracted for each study. Strong performing elements for each study are shown in shaded cells.

	Pizza SELF	Total	MS 1	MS 2	MS 3
	Base Size	30	15	7	8
	Additive Constant	62	52	81	71
	Strong for Mind-Set 1				
C2	We support community organizations and local charities	12	19	4	2
D2	Waters, sodas, and beverages available	4	19		
D1	Not just pizza - entrees, salads, wedges, appetizers ... and more	9	18		
D3	Catering ... any occasion ... any number ... any location	6	18		
C4	Dine-in ... table service available	10	16	7	6
C3	Warm, friendly staff treats you right	8	15		5
D4	Choose a specialty pie ... or order one exactly as you want it	4	15		
C1	Family owned, family traditions, family recipes	6	12		
	Strong for Mind-Set 2				
B1	Only the freshest, highest quality ingredients			18	
B4	Organic or natural ingredients ... used exclusively			13	
B2	Dough and sauce made fresh daily			9	
A2	Get deals and discount coupons ... from our site, app, or mailer	2		8	4
	Strong for Mind-Set 3				
A1	Order and pay online ... easy, convenient				24
A4	Pick-up Only				20
A3	Delivery available				13
	Not strong for any mind-set				
B3	Gluten-free pizza and pasta options			6	
	Pizza AI	Total	MS 1	MS 2	MS 3
	Base Size	31	11	10	10
	Additive Constant	44	19	52	60
	Strong for Mind-Set 1				
A4	Pineapple Pizza	4	36		
A3	Mushroom Pizza	2	29		
A2	Sausage Pizza		24		
A1	Cheese Pizza	7	20	9	
D1	Delivery is available, but I need to order a minimum amount.	8	18	4	
D4	Delivery is not available at this time.		15		
D3	Delivery is available from the pizzeria.	4	14	7	

D2	Delivery is available, but there is a delivery fee.		9		
	Strong for Mind-Set 2				
C1	The staff is friendly and welcoming.	4		20	
C4	The pizzeria has a TV playing in the corner.	1		20	
C2	The pizzeria has a family-friendly atmosphere.			15	
C3	The pizzeria is modern and trendy.			11	
A1	Cheese Pizza	7	20	9	
	Strong for Mind-Set 3				
B3	You may be able to find deals or specials if you are a member of a store's rewards program.	5			22
B4	Some stores offer deals or specials if you follow them on social media.	3	2		17
	Not strong for any mind-set				
B1	There may be deals or specials offered depending on the time of year and the store's inventory.				5
B2	You may be able to find coupons for deals or specials in newspapers, magazines, or online.	4		7	4
	Dry Cleaner - SELF	Total	MS 1	MS 2	MS 3
	Base Size	30	16	8	6
	Additive Constant	53	46	58	72
	Strong for Mind-Set 1				
B2	Premium care service available ... inspection, repair, hand finishing, packaging	9	18	7	
A4	Some services can be performed in home. furnishings cleaning, tailoring,	5	14		
B1	Wedding gown preservation	6	13	2	
A1	Full range of dry-cleaning services	10	12	10	10
A3	Laundry services ... wash, dry fold,	5	12		6
A2	Tailoring, hemming, sewing and alteration services	7	10		8
B3	Garment restoration ... from damage cause by fire, water, soot, or mold		8		
	Strong for Mind-Set 2				
D1	Free pick-up and delivery	4		20	2
D4	Coupons and discounts available ... through mailer's, website, app	2		17	2
D3	Convenient app ... arrange pickup, cleaning, delivery	6		16	9
C2	Focus on quality, value, service, and customer satisfaction	1	6	12	
A1	Full range of dry-cleaning services	10	12	10	10
C3	Dedicated and reliable professionals	6		10	15
	Strong for Mind-Set 3				
C3	Dedicated and reliable professionals	6		10	15
D2	Same day or next-day service				13
A1	Full range of dry-cleaning services	10	12	10	10
D3	Convenient app ... arrange pickup, cleaning, delivery	6		16	9
A2	Tailoring, hemming, sewing and alteration services	7	10		8
	Not strong for any mind-set				
B4	The right cleaning method properly matched to the job	1	5		
C1	Green cleaning and sustainable practices used ... minimize environmental impact	1		7	
C4	Multiple locations				4
	Dry Cleaner - AI	Total	MS 1	MS 2	MS 3
	Base Size	30	9	6	15
	Additive Constant	63	69	54	70

	Strong for Mind-Set 1				
A1	Call the company and schedule a pick-up or delivery.	3	19		3
D4	Seamstress on staff	5	14		11
A3	Contact the company through social media to schedule a pick-up or delivery.	4	9		6
	Strong for Mind-Set 2				
B3	We have a 100% satisfaction guarantee.	3		32	2
B4	We have the most experience.			21	
B2	We offer a loyalty rewards program.			16	
B1	We have the most 5-star reviews.	1		14	
C4	Consider the type of dirt or stain you are trying to remove.		6	14	
C3	Select a cleaning method that will not damage the surface you are cleaning.		3	12	
	Strong for Mind-Set 3				
A4	Leave a voicemail for the company to schedule a pick-up or delivery.				17
D4	Seamstress on staff	5	14		11
A2	Use the company's mobile app to schedule a pick-up or delivery.	5		1	8
	Not strong for any mind-set				
C1	Choose a cleaning method that is appropriate for the type of surface you are cleaning.		4	7	
C2	Choose a cleaning method that is safe for both you and the environment.		7	1	
D1	Fabric cleaning services				
D2	Customized services		5		1
D3	Rush orders	1	2		6
	Dentist SELF	Total	MS 1	MS 2	MS 3
	Base Size	31	10	9	12
	Additive Constant	51	44	41	49
	Strong for Mind-Set 1				
C2	The latest techniques and materials are used	5	21	14	
C3	The dentist is skilled and professional, but not very warm personally		18	8	
C1	Inviting office with friendly, supportive staff	4	15	19	
D1	You're provided with a personalized treatment plan	6	13	14	1
D4	The dentist takes a lifestyle approach to oral health, not just drilling and filling		10	1	
D3	No insurance ... no problem. Financing options available		9	1	
D2	Your insurance is accepted	1	8		
	Strong for Mind-Set 2				
C1	Inviting office with friendly, supportive staff	4	15	19	
A1	General Dentistry for the entire family	5		16	19
C2	The latest techniques and materials are used	5	21	14	
D1	You're provided with a personalized treatment plan	6	13	14	1
A3	Pediatric Dentistry and Orthodontics for kids			11	10
C3	The dentist is skilled and professional, but not very warm personally		18	8	
C4	Patients have an app for scheduling, remote visits, and the like		6	8	
	Strong for Mind-Set 3				
A1	General Dentistry for the entire family	5		16	19
A2	Cosmetic and Aesthetic Dentistry for a beautiful smile	2		1	15
A4	Advanced specialties like Endodontics and Oral Surgery for complex cases			5	15
B2	Your friends recommended this dentist	2	6		14

A3	Pediatric Dentistry and Orthodontics for kids			11	10
B3	You learned of the dentist through advertising		5		8
	Not strong for any mind-set				
B1	A dentist found by searching an online directory		6		1
	Dentist AI	Total	MS 1	MS 2	MS 3
	Base Size	30	12	8	10
	Additive Constant	66	83	46	66
	Strong for Mind-Set 1				
	Strong for Mind-Set 2				
D3	Dentist is able to provide you with financing options.	11	4	40	
D1	Dentist is able to provide you with a treatment plan.	7		38	
D4	Dentist is board certified.	10		37	
D2	Dentist is able to accept your insurance.	4	5	30	
A4	Orthodontics			12	
A1	General Dentistry			9	
	Strong for Mind-Set 3				
B2	The dentist is much better at explaining procedures than others in the area.	4			18
C2	Patients feel that their doctors are knowledgeable and up to date on the latest treatments.	4	5		13
C3	Patients feel that their doctors are accessible when they need them.	7	7	1	13
B4	The dentist is much more experienced than others in the area.				12
C1	The patient experience is generally positive, with patients feeling cared for and well-informed.	1	3		9
B1	The dentist is much more gentle when cleaning teeth than others in the area.	5	3	1	8
	Not strong for any mind-set				
A2	Cosmetic Dentistry				
A3	Pediatric Dentistry				7
B3	The dentist is much better at making patients feel comfortable than others in the area.	1		5	6
C4	Overall, patients are satisfied with their doctor and the care they receive.				
	School Board SELF	Total	MS 1	MS 2	MS 3
	Base Size	30	8	11	11
	Additive Constant	58	77	33	69
	Strong for Mind-Set 1				
D2	Adopt "healthy start times" ... Start the day later for high schoolers	4	12		
C4	Increase student safety and security in our schools	9	8	27	
	Strong for Mind-Set 2				
C4	Increase student safety and security in our schools	9	8	27	
C3	Stop administrators from micro-managing our schools	2		23	
C1	End closed-door decision-making ... make decisions with the advice of the community and teachers	6	6	22	
C2	Make sure our technology and physical spaces inspire learning	2		18	
B3	Work with state and local representatives to increase school budgets	9		16	28
A2	Prepare each student to succeed in whatever future awaits them ... regardless of their background or circumstances			15	
A1	Close the achievement gap from COVID			11	
	Strong for Mind-Set 3				
B2	Reduce waste in school spending ... scrutinize every line item	1			30
B3	Work with state and local representatives to increase school budgets	9		16	28

B4	Seek school construction grants ... upgrade or replace schools	2			19
B1	Apply for grants that supplement the school budget	4			17
	Not strong for any mind-set				
A3	Kids should learn in school ... use remote learning only when necessary			7	
A4	Open more classrooms and reduce class sizes			3	
D1	Make it easier to get to and from school ... more efficient bus runs, better sidewalks and crossings, easier pickup and drop off				
D3	Create an orientation center for immigrant students before placing them in the classroom			4	
D4	Balance academics with emotional and mental wellness				
	School Board AI	Total	MS 1	MS 2	MS 3
	Base Size	30	2	15	13
	Additive Constant	60	100	74	34
	Strong for Mind-Set 1				
	Strong for Mind-Set 2				
	Strong for Mind-Set 3				
B4	Provide targeted support to students who are struggling academically	1			22
B1	Make schools more culturally relevant and responsive to the needs of students				15
C2	Encourage students to take challenging coursework.	1			14
B3	Use data to identify and target areas of need within the school system				13
B2	Teach social and emotional skills along with academics	1			12
A1	Improve security at school entrances and exits				11
	Not strong for any mind-set				
A2	Review and update safety protocols on a regular basis	2			5
A3	Foster a culture of safety and vigilance				4
A4	Keep doors and windows locked at all times				4
C1	Teach students study habits and time management skills.				
C3	Help students develop critical thinking skills.				
C4	Help students set goals for themselves and develop a plan to achieve those goals.				4
D1	We can address the issue of school funding by creating a new funding stream.				
D2	We can address the issue of school funding by redirecting existing funds.	5		5	7
D3	We can address the issue of school funding by cutting spending in other areas.			3	
D4	We can address the issue of school funding by increasing fees and charges.				2

Each respondent was a member of an online panel, aggregated by Luc.id Inc. The respondents were invite to participate by an email sent only to panel members. The panels comprised more than several hundred thousand individuals for each panel company whose members were aggregated by Luc.id, making recruiting easy. Those respondents who agreed to participate read an introductory statement, completed a short self-profiling questionnaire, and evaluated a unique set of 24 vignettes, created according to an underlying experimental design (Gofman & Moskowitz, 2010). The respondents rated the vignette on an anchored 5-point scale, shown below.

The screen below describes a local pizzeria. Please indicate how likely you would be to patronize this pizzeria on a scale of 1-5. Although some screens may look alike, each one is different. Don't

dwell on them, just choose a number based on your gut reaction. 1=Not at all likely ... 5=Extremely likely.

For each dentist described on this screen, please rate how likely you would be to choose this dentist. 1=Not at all likely ... 5=Extremely likely.

You will be shown a series of screens. Each one describes a local dry cleaner. For each dry cleaner described on a screen, please rate how likely you would be to patronize this dry cleaner.

1=Not at all likely ... 5=Extremely likely.

Based on the policy positions of the school board candidate below, how likely would you be to vote for this person? 1=Unlikely ... 5=Extremely likely.

The analysis converted ratings 5 and 4 to 100, ratings 1-3 to 0, added a vanishingly small random number to each transformed number, used OLS (ordinary least-squares regression), and then clustered the respondents into two and then three mind-sets using k-means clustering. The OLS regression related the presence/absence of the 16 elements to the transformed rating scale, doing so at the level of each respondent. Each respondent ended up with an additive constant and 16 coefficients. The additive constant shows the basic level of interest in the topic (e.g., voting for a person for the school board) in the absence of elements, value that ends up becoming a baseline value since all vignettes comprised 2-4 elements (viz. answers) as dictated by the underlying experimental design.

The respondents were then divided first into two non-overlapping but exhaustive groups, and then into three non-overlapping but exhaustive groups based upon K-Means Clustering. The 16 coefficients were used as the basis for clustering, with the additive constant ignored for by the clustering algorithm.

Table 1 below divides into eight sections, one section for each of four topics, done twice. The two ways are done by the researcher alone, SELF using the AI Coach, the second way, AI, done by the researcher using the AI-powered Idea Coach. Table 1 shows the results for Total Panel and for three mind-sets, for each pair of studies One can get a sense of the richness of the ideas generated by the researcher (SELF) and generated by the researcher aided by AI in the form of Idea Coach.

A cursory look at Table 1 suggests similar performances by elements chosen by a person, first without AI (SELF), and later, at a separate time, with the help of AI (AI). There are no glaring patterns which emerge to tell us that AI-augmented efforts are dramatically stronger or weaker, when selected for the same topic, and evaluated by individuals are presented with a test protocol which prevents guessing what the underlying patterns might be. Only in one of the four studies do we see clear superiority, with elements for school board chosen by a person element performing far better than elements chosen by a person selecting from offerings of AI. This means AI does not offer up any better ideas than the ideas emerging from the individual himself.

Creating a Metric for Comparison – IDT (Index of Divergent Thinking)

A different metric is called for, to compare the performance of the study. This metric has been used for Mind Genomics studies. Table 2 presents the computation for the IDT, index of divergent thinking. In simple terms, the IDT looks at the weighted sum of positive coefficients

generated by the combination of the three major groups: total panel, both mind-sets in the two-mindset solution, all three mind-sets in the three mind-set solution. Each of these three components contributes an equal proportion to the final IDT. The approach makes sense because it is very difficult to generate a high coefficient for the total panel because the different people in the total panel cancel each other out. Thus, the total panel gets a weight of 33%. When it comes to two mind-sets, they also share a weight of 33%, so the two mind-sets compete for the 33%. Finally, when I come to three mind-sets, they also share a weight of 33%, so the three mind-sets compare for their 33%. Table 2 shows the computation.

When the IDT is computed for the eight studies, the results give a sense of the subjective ‘strength’ of the elements, when ‘strength’ is operationally defined as strong performance on the rating attribute (viz., a preponderance of assigned ratings of 5 and 4 to the vignettes by a defined subgroup of respondents). Table 3 presents a comparison among the eight studies, the sum of positive coefficients for each of the key groups (total, two mind-sets, three mind-sets, respectively), then the IDT for the study.

The important number in Table 3 is the IDT for virtually the same number of respondents across four pairs of matched studies, one study with four questions and four sets of four elements chosen completely by the researcher without any help from AI, and then afterwards the same selection, this time using only the questions and answers provided by the selected OpenAI language model and based only on a simple statement about the topic.

Table 3: Key statistics from the eight studies.

Topic	Method	Sum of all positive coefficients for the subgroup						IDT
		Total	MS 1 of 2	MS 2 of 2	MS 1 of 3	MS 2 of 3	MS 3 of 3	
Dry Cleaner	AI	22	50	42	67	69	118	54
	Self	63	79	85	76	98	101	69
Dentist	AI	54	70	43	126	27	173	86
	Self	25	69	103	64	120	98	83
Pizza	AI	42	79	118	58	167	93	48
	Self	61	86	132	40	132	65	74
School Board	AI	10	25	0	14	0	8	113
	Self	39	63	25	78	26	146	94

Table 2: Computation of the IDT, the ‘Index of Divergent Thought’. The IDT provides a metric for the strength of the ideas using subjective judgments of the vignettes.

Pizza - SELF	Total	MS 1 of 2	MS 2 of 2	MS 1 of 3	MS 2 of 3	MS 3 of 3
A Sum of All Positive Coefficients	61	132	40	132	65	74
B Average of All Positive Coefficients (A/C)	2.0	7.3	3.3	8.8	9.3	9.3
C Number of respondents in the column (subgroup)	30	18	12	15	7	8
D Weight = % respondents in subgroup (C/Total C)	0.33	0.20	0.13	0.17	0.08	0.09
E Weighted Total (Sum x Weight; A x D)	20.1	26.4	5.2	22.4	5.2	6.7
F IDT Final Score = sum of weighted totals (Total E)	86.0					

At this stage, we can say that the results are inconclusive. AI generates a much higher IDT value for school board (113 vs. 94, AI vs. Self), similar value for dentist (86 vs. 83), lower for dry cleaner (54 vs. 69) and much lower for pizza shop (48 vs. 74).

These initial results suggest that AI may perform better than Self generates messages for topics which are not ordinary, not daily (viz., dentist, and school board), but will probably perform more poorly for the more common topics (viz., dry cleaner, pizza shop). The respondents were matched in each study in terms of market and age, and the study was not identified as to origin of the elements. It may be that the differences could emerge from discrepancies in the base sizes of the two groups of emergent mind-sets (MS1 of 2 and MS2 of 2 vs. MS1 of 3, MS2 of 3, and MS3 of 3). If we just look at the Total Panel, using the Total of the positive coefficients, we find that only dentist data shows a stronger performance of AI generated elements versus self-generated elements.

Are AI Generated Elements More or Less ‘Engaging’ than Self-generated Elements?

Up to now the data strongly suggest that the current AI generated elements do not perform quite as well as the self-generated elements, although they do not perform poorly, at least when the judgment is cognitive. What happens, however when the metric comparing the two is non-cognitive, and often used by consumer researcher as well

as experimenters in psychology and other disciplines’ course, we refer here to response time, a time-honored method in psychology and more recently in consumer research (Bassili & Fletcher, 1991).

The Mind Genomics program, BimiLeap, measures the time between the presentation of the test stimulus, the vignette, and the response to the stimulus, viz., the rating assigned by the respondent. The respondent need not do anything but read and respond. The time between the presentation of the vignette and the respondent, response time, become the dependent variable in a simple regression equation:

$$RT = k_1 A1 + k_2 A2 \dots k_{16} D4$$

The response time model does not have an additive constant, simply because the additive constant does not have any real meaning. The dependent variable, response time, is used in place of the 5-point scale, either at the level of the group, or for other studies, at the level of the individual respondent.

Table 4 shows the response time estimate for each element for all eight studies. The table shows the longer response times in shade. This paper does not evaluate the response time, but two opposite arguments can be made about the response time. The first argument is that the longer response times ‘engage’ the respondent to pay attention. The second argument is that the longer response times due to the fact that the message may take longer to read because the message is confusing, poorly written, or simply has mor words.

Table 4: Estimated response time for each element, estimated by Mind-Genomics.

	SELF		AI (Idea Coach)	
Pizza Shop				
	Question A: Price, Ordering & Delivery		Question A: What type of pizzas will be offered?	
A1	Order and pay online ... easy, convenient	0.9	Cheese Pizza	0.8
A2	Get deals and discount coupons ... from our site, app, or mailer	0.8	Sausage Pizza	1.0
A3	Delivery available	0.8	Mushroom Pizza	0.8
A4	Pick-up Only	0.9	Pineapple Pizza	0.9
	Question B: Ingredients		Question B: Will there be any deals or specials?	
B1	Only the freshest, highest quality ingredients	0.9	There may be deals or specials offered depending on the time of year and the store's inventory.	1.0
B2	Dough and sauce made fresh daily	0.7	You may be able to find coupons for deals or specials in newspapers, magazines, or online.	1.1
B3	Gluten-free pizza and pasta options	1.1	You may be able to find deals or specials if you are a member of a store's rewards program.	1.2
B4	Organic or natural ingredients ... used exclusively	0.9	Some stores offer deals or specials if you follow them on social media.	1.2
	Question C: Culture		Question C: What is the atmosphere of the pizzeria?	
C1	Family owned, family traditions, family recipes	0.9	The staff is friendly and welcoming.	0.8
C2	We support community organizations and local charities	0.5	The pizzeria has a family-friendly atmosphere.	0.7
C3	Warm, friendly staff treats you right	0.8	The pizzeria is modern and trendy.	0.9
C4	Dine-in ... table service available	0.8	The pizzeria has a TV playing in the corner.	0.9
	Question D: Menu & Offerings		Question D: Does the pizzeria offer delivery?	
D1	Not just pizza - entrees, salads, wedges, appetizers ... and more	0.9	Delivery is available, but I need to order a minimum amount.	0.8
D2	Waters, sodas, and beverages available	0.8	Delivery is available, but there is a delivery fee.	0.7
D3	Catering ... any occasion ... any number ... any location	1.0	Delivery is available from the pizzeria.	1.1
D4	Choose a specialty pie ... or order one exactly as you want it	1.1	Delivery is not available at this time.	1.0

Local Cleaner				
	Question A: Typical Services Offered		Question A: How do I schedule a pick-up or delivery?	
A1	Full range of dry cleaning services	0.5	Call the company and schedule a pick-up or delivery.	1.0
A2	Tailoring, hemming, sewing and alteration services	0.4	Use the company's mobile app to schedule a pick-up or delivery.	1.0
A3	Laundry services ... wash, dry fold,	0.4	Contact the company through social media to schedule a pick-up or delivery.	0.9
A4	Some services can be performed in home .. furnishings cleaning, tailoring,	0.3	Leave a voicemail for the company to schedule a pick-up or delivery.	1.0
	Question B: High-end Services		Question B: What sets your business apart from other dry cleaners?	
B1	Wedding gown preservation	0.9	We have the most 5-star reviews.	1.0
B2	Premium care service available ... inspection, repair, hand finishing, packaging	0.9	We offer a loyalty rewards program.	0.9
B3	Garment restoration ... from damage cause by fire, water, soot or mold	1.1	We have a 100% satisfaction guarantee.	1.1
B4	The right cleaning method properly matched to the job	0.8	We have the most experience.	1.1
	Question C: Professionalism		Question C: How do you decide which cleaning methods to use?	
C1	Green cleaning and sustainable practices used ... minimize environmental impact	0.9	Choose a cleaning method that is appropriate for the type of surface you are cleaning.	0.3
C2	Focus on quality, value, service and customer satisfaction	0.6	Choose a cleaning method that is safe for both you and the environment.	0.3
C3	Dedicated and reliable professionals	0.7	Select a cleaning method that will not damage the surface you are cleaning.	0.7
C4	Multiple locations	0.7	Consider the type of dirt or stain you are trying to remove.	0.3
	Question D: Convenience and Price		Question D: Do you offer any special services like in-home alteration or fabric cleaning?	
D1	Free pick-up and delivery	0.9	Fabric cleaning services	0.7
D2	Same day or next-day service	0.6	Customized services	0.7
D3	Convenient app ... arrange pickup, cleaning, delivery	1.0	Rush orders	0.7
D4	Coupons and discounts available ... through mailer's, website, app	0.7	Seamstress on staff	0.8
Dentist				
	Question A: What is the range of dental services offered?		Question A: What is the range of dental services offered?	
A1	General Dentistry for the entire family	0.9	General Dentistry	0.9
A2	Cosmetic and Aesthetic Dentistry for a beautiful smile	0.8	Cosmetic Dentistry	0.7
A3	Pediatric Dentistry and Orthodontics for kids	1.1	Pediatric Dentistry	0.8
A4	Advanced specialties like Endodontics and Oral Surgery for complex cases	1.2	Orthodontics	1.0
	Question B: Which sources of information do you use?		Question B: How does the dentist compare to others in the area?	
B1	A dentist found by searching an online directory	1.3	The dentist is much more gentle when cleaning teeth than others in the area.	0.7
B2	Your friends recommended this dentist	1.3	The dentist is much better at explaining procedures than others in the area.	0.7
B3	You learned of the dentist through advertising	1.4	The dentist is much better at making patients feel comfortable than others in the area.	0.7
B4	The dentist has a great social media presence	1.4	The dentist is much more experienced than others in the area.	0.5
	Question C: What is the office experience like?		Question C: What is the patient experience like?	
C1	Inviting office with friendly, supportive staff	1.2	The patient experience is generally positive, with patients feeling cared for and well-informed.	0.9
C2	The latest techniques and materials are used	1.2	Patients feel that their doctors are knowledgeable and up to date on the latest treatments.	0.9
C3	The dentist is skilled and professional, but not very warm personally	1.4	Patients feel that their doctors are accessible when they need them.	1.0
C4	Patients have an app for scheduling, remote visits, and the like	1.1	Overall, patients are satisfied with their doctor and the care they receive.	0.9
	Question D: What is the treatment approach and insurance options?		Question D: What qualifications should I look for in a dentist?	
D1	You're provided with a personalized treatment plan	1.3	Dentist is able to provide you with a treatment plan.	0.6
D2	Your insurance is accepted	1.2	Dentist is able to accept your insurance.	0.7
D3	No insurance ... no problem. Financing options available	1.2	Dentist is able to provide you with financing options.	0.6
D4	The dentist takes a lifestyle approach to oral health, not just drilling and filling	1.5	Dentist is board certified.	0.8

School Board				
	Question A: How can we improve educational outcomes?		Question A: How can we address the issue of school safety?	
A1	Close the achievement gap from COVID	0.9	Improve security at school entrances and exits	0.9
A2	Prepare each student to succeed in whatever future awaits them ... regardless of their background or circumstances	0.7	Review and update safety protocols on a regular basis	0.7
A3	Kids should learn in school ... use remote learning only when necessary	0.8	Foster a culture of safety and vigilance	0.8
A4	Open more classrooms and reduce class sizes	0.7	Keep doors and windows locked at all times	1.1
	Question B: How can we improve the financial stability of our schools?		Question B: What can be done to address the achievement gap in our schools?	
B1	Apply for grants that supplement the school budget	0.9	Make schools more culturally relevant and responsive to the needs of students	1.0
B2	Reduce waste in school spending ... scrutinize every line item	1.0	Teach social and emotional skills along with academics	0.9
B3	Work with state and local representatives to increase school budgets	1.3	Use data to identify and target areas of need within the school system	1.1
B4	Seek school construction grants ... upgrade or replace schools	1.1	Provide targeted support to students who are struggling academically	1.0
	Question C: How can we improve security, the school board, integrate immigrants, and run schools better?		Question C: How can we better prepare our students for college and careers?	
C1	End closed-door decision-making ... make decisions with the advice of the community and teachers	1.1	Teach students study habits and time management skills.	1.0
C2	Make sure our technology and physical spaces inspire learning	0.9	Encourage students to take challenging coursework.	0.8
C3	Stop administrators from micro-managing our schools	1.2	Help students develop critical thinking skills.	1.0
C4	Increase student safety and security in our schools	1.4	Help students set goals for themselves and develop a plan to achieve those goals.	0.9
	Question D: How can we improve transportation and support our students' growth		Question D: How can we address the issue of school funding?	
D1	Make it easier to get to and from school ... more efficient bus runs, better sidewalks and crossings, easier pickup and drop off	0.8	We can address the issue of school funding by creating a new funding stream.	0.9
D2	Adopt "healthy start times" ... Start the day later for high schoolers	0.8	We can address the issue of school funding by redirecting existing funds.	1.0
D3	Create an orientation center for immigrant students before placing them in the classroom	0.6	We can address the issue of school funding by cutting spending in other areas.	0.9
D4	Balance academics with emotional and mental wellness	0.8	We can address the issue of school funding by increasing fees and charges.	1.0

Discussions and Conclusions

The rapid emergence of 'cognitively rich' AI cannot be ignored. Newspaper articles discussing trends of AI point to the ability of widely available AI tools to create prose in a way which mimics that prose written by people who are capable writers, producing what could be called felicitous prose. Indeed, the topic of the growing potential of students to write college essays using AI is forcing a reevaluation of what it means for a student to learn to write, or indeed to get a liberal arts education. When the language generated by AI is sufficiently close to the language 'ordinary people' use, a new paradigm in called for in education.

The contribution of this paper is to introduce the human judge into the evaluation of snippets of information, viz., the elements produced either by people or by artificial intelligence. The Mind Genomics study does not focus on the coherence of the composition, viz., the 'fitting together' of the elements into a coordinated paragraph. Rather, the Mind Genomics effort in this paper is to understand the degree to which respondents feel about the individual texts, and to determine whether the elements generated by AI produces the same evaluation score of feeling as does the text of the same topic produced by a person.

If one were to summarize the findings of this paper, one would have to conclude that in terms of human judgment, the elements generated

by AI may approach the quality of the elements generated by one's mind alone, but only in one study out of four (school board) do we see AI performing better, presumably because people do not know about school boards in the way they know about pizza shops, dry cleaners, and dentists. Furthermore, the response time for elements are higher for elements created by the person, rather than by AI. This might be because the text elements created by the highly regarded AI used for the study (Zuccarelli 2020) were sensible, but bland and simply not as engaging yet as those written by a person.

References

- Bassili JN, Fletcher JF (1991) Response-time measurement in survey research a method for CATI and a new look at nonattitudes. *Public Opinion Quarterly* 55: 331-346. [\[crossref\]](#)
- Dedehayir O, Steinert M (2016) The hype cycle model: A review and future directions. *Technological Forecasting and Social Change* 108: 28-41. [\[crossref\]](#)
- Fyfe P (2022) How to cheat on your final paper: Assigning AI for student writing. *AI & SOCIETY*, pp.1-11. [\[crossref\]](#)
- Gofman A, Moskowitz H (2010) Isomorphic permuted experimental designs and their application in conjoint analysis. *Journal of Sensory Studies* 25: 127-145. [\[crossref\]](#)
- Grace K, Salvatier J, Dafoe A, Zhang B, Evans O (2018) When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research* 62: 729-754. [\[crossref\]](#)
- Hill J, Ford WR, Ferreras IG (2015) Real conversations with artificial intelligence: A comparison between human-human online conversations and human-chatbot conversations. *Computers in Human Behavior* 49: 245-250. [\[crossref\]](#)

7. 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. [\[crossref\]](#)
8. Zuccarelli, Eugenio (2020) GPT-3, OpenAI's Revolution: Looking into the most powerful language model ever created, Towards Data Science, GPT-3, OpenAI's Revolution. Looking into the implications of the... | by Eugenio Zuccarelli | Towards Data Science, accessed 20 December 2022 [\[crossref\]](#)

Citation:

Moskowitz H, Rappaport S, Deitel Y (2023) The Quality of Ideas When AI (Artificial Intelligence) is Used as a Coaching Device. *Ageing Sci Ment Health Stud* Volume 7(2): 1-12.