ATTICUS ENTRY: “The Real Why & The Hidden Who”
CATEGORY: Data-driven Insights
AUTHORS: Christopher Graves (Founder and President, Ogilvy Center for Behavioral Science) and Jon Puleston (VP Innovation, Kantar Profiles).

This entry won the ESOMAR World Congress top prize for best paper on September 11, 2019.

OVERVIEW
Behavioral scientists have developed a variety of tools to test and diagnose individuals in order to better understand and predict behavior, choices and preferences. These include: personality trait profiles; the cultural cognition (worldviews) that reinforce a group’s undisclosed identity and values and guide opinions despite facts; and how people are wired to make sense of communications (cognitive styles). But these many varied tools were used in isolation, and not integrated into one sort of decision-making genome. And they had not been used in the realm of marketing and communications. Frankly, behavioral science has started to prove that much of the established, conventional wisdom of agency-led, behavior-change thinking has been deeply flawed.

The opportunity to create a new communications-behavioral genome led to a multi-year collaboration between Christopher Graves (Ogilvy) and Jon Puleston (Kantar). Graves captured more than 10,000 scientific studies, and created a new marketing-friendly taxonomy and app to sort and collate them. He unearthed very robust—but completely unknown in the world of marketing and communications—scientific methods of decoding individuals to better decipher and predict their preferences and behavior. But these many tests had never been integrated into one practical, commercial instrument at scale. Jon Puleston (Kantar) created a new integrated instrument, while improving the underlying behavioral science tools and improving the test methodologies. His methods reduced self-report bias, a possible weakness of the original scientific tools. And his large-scale field testing of the work-in-progress led to a better and better research instrument.

The result—an award-winning, breakthrough approach to decoding the “Real Why” of human behavior (as opposed to what people may say or what conventional wisdom may wrongly assume), and the “Hidden Who” of individual makeup that reveals a predictive blueprint to craft more effective creative, content and communications tailored to individual decision-making genomes. Since the invention, the method is being deployed on some of the toughest human challenges, from vaccine hesitancy, to smoking cessation, to communicating climate change, and even fighting extremist recruiting.

Below is the presentation delivered live by Graves & Puleston at the ESOMAR World Congress, and their ESOMAR Award-winning paper.
The Real Why & The Hidden Who
By Christopher Graves (Ogilvy) and Jon Puleston (Kantar)

THIS ATTICUS ENTRY CONSISTS OF TWO PARTS:

Pages 03-19  PART 1: the ESOMAR Congress speech & presentation live Sept. 9, 2019

Pages 21-42  PART 2: the ESOMAR Congress “Best Paper” Winner  (Sept. 11, 2019)
ATTICUS ENTRY PART 1: “THE REAL WHY & THE HIDDEN WHO” SPEECH

By Christopher Graves (Ogilvy) and Jon Puleston (Kantar)

ESOMAR CONGRESS SPEECH (“The Real Why & The Hidden Who”)

The document that follows is the presentation given live at the ESOMAR World Congress in Edinburgh on September 9, 2019 that is the companion piece to our award-winning research paper (Part 2).
Early in his career, David Ogilvy worked as a researcher. He has been quoted as complaining that a shortcoming in market research is that people “don’t say what they think and don’t do what they say.”

The past couple of decades of emerging behavioral science have proven Ogilvy right. When asked, people will tell you why they did something or will do something—but it is unlikely they themselves know why. Using new findings in behavioral science can help us decode the “Real Why” of human behavior.

When we say “behavioral science,” we mean social psychology, cognitive neuroscience, evolutionary psychology and behavioral economics. Two Nobel Prizes have been won for behavioral economics—one just this past year. And a number of books have begun to broaden our understanding of behavioral economics. They tell us about the so-called “cognitive biases” hardwired into our 40,000 year-old brains that lead us to quick but sometimes inaccurate decision-making, and how we may “Nudge” people into better decisions and healthier behavior.

But we have gone beyond the nudge. We have integrated, for the very first time, a battery of robust, behavioral science instruments, that together, can reveal an individual’s decision-making genome.
We have digested more than 10,000 studies over the past decade and remapped them against the challenges marketers and communicators face every day. We’ve divided the findings along two tiers: the high-level of our human species tier where our anachronistic human software creates these cognitive biases that affect us all to a greater or lesser degree. Things like why we fear shark bites more than what will really kill us—heart disease. Or why we don’t really care about the future (despite professing otherwise) when it comes to climate change personal action or saving for retirement. Or why we take our behavioral cues from others no matter how independent-minded we think we are. Or why we resist the miracle of vaccinations. Or why warning labels don’t make people quit smoking, nor stop eating foods that make them obese. Or why consumer decisions can seem baffling, unpredictable, despite traditional consumer research.

As the behavioral scientist and author Jonathan Haidt puts it: “Reason is the press secretary for the emotions.”

We also work at decoding humans at the individual level. We have integrated three lenses taken from deep, robust scientific studies that have not previously been used together, and never used for marketing and communications.

Using these lenses we have found we can reveal the “Hidden Who”—the individual decision-making genome that filters, clouds and guides all human choices, preferences and behavior.
The first lens is personality trait science. This is not some flimsy fad, nor is it Myers-Briggs, which though widely used is not deemed a real science by the scientific community. But personality trait science has been confirmed in thousands of studies over the past 25 years.

The most widely accepted version is called “Big Five,” or “Five Factor” OCEAN, named for the five factors that make up a person’s personality. Each factor is on a spectrum from low to high. Find yourself here with our celebrity personality test. Openness to Experience runs from a low end comfort with routine and tradition (Prince Phillip) to a high end explorer craving variety (Richard Branson). Low Conscientiousness - think Ricky Gervais, and on the high end Hermione Granger. High Extraverts love being with others and in the limelight (Graham Norton) vs introverts like Bill Gates. Someone low on Agreeableness seems disbelieving, skeptical maybe cranky and a bit more self-interested (Larry David) vs someone on the high end who gushes warmth and likability (Jennifer Anniston). And finally Neuroticism. This does not mean you are clinically neurotic—it is more a worry and anxiety threshold. Those who worry easily (Woody Allen) score higher.

You can read more about this and how we can use personality trait science in marketing in my Harvard Business Review article called “What Marketers Should Know About Personality-Based Marketing”.

---

**BIG 5 “OCEAN”**

- **Openness**
- **Conscientiousness**
- **Extraversion**
- **Agreeableness**
- **Neuroticism**

---

**What Marketers Should Know About Personality-Based Marketing**

---

**WORLDVIEW**

Cultural Cognition
Our second lens is Cultural Cognition—or your Worldview. To illustrate how this invisible force field plays out in pop culture, look at this heat map of TV viewership of two popular US shows. “Modern Family” features all sorts of inclusivity—age, gender, LGBTQ, race… while “Duck Dynasty” features four isolationist swamp dwellers. They are negative images of each other.

Cultural Cognition is a powerful and robust framework born of the work of scientist Dan Kahan at Yale University. It can map individuals, groups or even brands across two axes. Atop the vertical axis sit those who believe the world works best when, in every situation, there is a strong leader and a clear hierarchy below them. Their opposites, Egalitarians, gravitate toward group and consensus rather than hierarchical social structures. On the horizontal axis at left sit Individualists. They believe all accountability and responsibility for your personal outcomes rest solely with you—not fate, bad luck or government help—fix it yourself. Their opposites, Communitarians, believe we are all one human family and are obliged to close gaps of unfairness such as income, race, gender or other biases. The confluence of two forces forms a quadrant, and in each quadrant a clearer portrait of what motivates that individual or group. For example, Hierarchical-Individualists believe that free markets are the best solution to about any problem, that individual opportunities within law and order matter most. And they adhere to traditional, perhaps nostalgic societal norms. Their diagonal opposites want to protect the vulnerable while changing unfair, current societal ills.
Our third lens is actually a collection of “cognitive styles,” i.e. how we as individuals are wired to make sense of the world, of communications and of content.

Here are a few we use:

“Locus of Control” is a spectrum between external and internal. Someone testing for an external locus of control feels outside forces have more agency and control over their life. It could be fate, chance, religion, bureaucracies or government or indeed, their own genes. An internal LOC is like Steve Jobs. They make it happen and don’t care about outside rules, norms or barriers.

“Regulatory Focus” divides between those with more of a promotion ("go for it") mindset versus those with a prevention (risk no harm) mindset. People may test for a higher “Need for Affect” (moved by emotional stories) or “Need for Cognition” (need proof and like investigating more). Everyone defaults to a sort of “time perspective” in the those rooted in the present can be more hedonic, caring little for future consequences (they will struggle with diets, exercise plans and saving for retirement). Those more future-minded will more readily and conscientiously embrace behavior change today - resulting in sounder and healthier outcomes in the future.

All these tests overlap like lenses. They make the individual portrait increasingly clearer, from consumer or patient insights, creating better personas and more precise journey maps. Thus better fits are made between the individual decision-making genome and content and between creative execution and messaging.

HOWEVER, there was a problem with self-report bias in the methodologies. Some individuals struggle with objective self-awareness when filling out surveys (example at left) …
This chart illustrates the extent of the problem we face with classical personality testing techniques. If you ask people to self-assess their own personality nearly everyone places themselves into the top half of scales, measuring positive attributes like conscientiousness or openness. Few people directly admit to having any negative personality traits, for example of being critical or disorganized.

This leads to very homogenized, poorly differentiated data. When trying to use classical tests in commercial research projects we found that nearly everyone looks the same.

Furthermore, we found it very hard to compare results from one country to another because of the way different cultures answer the scale questions in classical tests; some even more positively than others.

The other challenge we faced, in particular when using the Big 5 personality test, was that these broad measures were often not quite nuanced enough to help adequately explain and reveal the whole story about the personality of a target audience.

This is exemplified by some work we undertook to understand the challenges of weight loss. A behavioral scientist had hypothesized that people struggling to lose weight were more open and had lower levels of conscientiousness. But when tested, their openness measure was
Together we set out to completely rethink how to more effectively measure personality in a way that would be more useable for commercial market research. Our aim was to be better able to differentiate audiences and to provide a more accurate viewpoint: by evaluating not just a persona personality but crucially, how they think and see the world.

We started by taking a critical look a methodology.

Nearly all personality measurement techniques are dominated by long drawn out, repetitive sets of Likert scale questions.

The issue with Likert scales, is that when people are unsure how to answer a question, and, as they get bored of answering questions, they have a strong tendency to slightly agree with everything! This results in a lot of noisy, inconsistent data and/or findings like the example here, where 30% of people agree that they are both chatty and quiet.

Our solution was to switch some of the measures to use a technique we have pioneered called choice-based prioritization. We present clusters of competing personality traits that we asked people to pick out and prioritize.

We have found this delivers much cleaner data with almost no overlap of conflicting choices.

Also, it has the added advantage of being quicker. We can evaluate up to 3 times as many personality dimensions in the same amount of time.

However, this technique goes only so far in solving the problem....
The method proved very successful at tackling the challenge of measuring positive personality measures. Faced with competing choices, people tend to only pick the descriptors that they most strongly identify with and this reduces the over claim.

What it did not solve was the challenge of measuring more negative personality traits. Few people will click on a button that says I am disorganized.

So, we realized we needed to segment out how we measured negative personality dimensions and to explore different approaches here.

In our quest to find a solution we tested out a range of methods. One of the most successful was to simply highlight to people the challenge of self-assessment, using the Grumpy Cat meme.

When humorously raising awareness that we don’t like to admit to certain things about ourselves, we found that people opened up and then gave us a lot more candid feedback.

But probably the most important breakthrough we made was a shift to using behavioral measures that people found easier to answer truly.

For example. If you ask people straight out, “Are you a disorganized person?” only 12% admit to being this. If however you show them 2 clothes drawers: one organized and one disorganized, more than half pick the disorganized one.

Now we realized that it would unfair to base an assessment of someone’s overall level of conscientiousness based solely on their clothes drawers, but it does provide a small clue.
We found that combining multiple behavioral measures like these could provide a much more reliable assessment of negative personality dimensions than traditional approaches.

In developing our new approach, we tested out literally hundreds of these different behavioral measures, alongside more classical questioning techniques. We then used principal component analysis to isolate the most useful questions to measure for each of the core personality dimensions.

We must stress that in the revised test we devised, we have not abandoned using Likert scales altogether, as they still proved very useful for measuring certain types of more neutral personality dimensions that were less affected by confirmation bias.

What we did was to introduce a wider range of question approaches, each specifically used to measure different aspects of someone’s personality.

This made the survey less repetitive and more interesting for participants, which in turn reduced the negative impact of respondent fatigue.
Having established a new framework for measuring personality, our next challenge was to hone down exactly what we wanted to measure.

Our goal was to devise - in one test - a means of measuring the whole person’s thinking processes.

We wanted to measure not just someone’s primary personality traits but some of the most important personality sub-dimensions that are most relevant to consumer decision making. For example: conscientiousness can be broken down into 3 sub dimensions: how careful are you and your attention to detail, your organizational/planning skills and how goal driven you are.

We wanted to add a second lens to further understand how a person sees the world (their worldview or “cultural cognition”), what their social priorities and important issues were. This approach was pioneered by the Cultural Cognition Project at Yale University.

A third lens was to measure how they make sense of things, to map out a person’s key cognitive thinking styles.

We isolated what we thought were the 7 most commercially useful measures of cognitive thinking styles: Regulatory Focus, to understand a person’s attitude towards risk; Locus of Control and self-efficacy to measure a person’s sense of empowerment; Time Perspective, to see if someone focuses on the future or is more present biased; Need for Affect and Need for Cognition to understand how a person engages with emotions and likes to think (useful for creative executions). Finally, an assessment of a person’s consumer motivations to understand how utilitarian & rational, or how emotional & hedonic their shopping decision-making processes are.
The challenge of measuring all these things in one go is one of optimization and economy. We whittled each into 2 or 3 minute modules, (a sketch rather than a detailed portrait).

We discovered however, that combining multiple measures into one survey offered significant cross validation benefits, akin to a Blockchain protocol.

Every different personality dimension is closely inter-woven. For example, our levels of conscientiousness are driven by our sense of self efficacy: how focused and how open we are.

Thus, we found the answers from each part of the test could be used to cross validate other parts.

Our core measure of conscientiousness is made up of 10 core questions but is validated by the answers to 20 other questions from other personality dimensions.

This leads to a lot more cross-cultural stability too, as it means that if one or two questions are not interpreted the same way in a particular country, it does not corrupt the whole measure.

The final important element to devising this test was to design the survey, which at around 17 minutes long required a lot of concentration from participants.

We paid a great deal of attention to the visualization of the survey experience to signpost content and topic shift.
At the start of each 3 minute section we used a “thought starter” (a fun question to signal the topic shift and to spark their imagination).

At the end of each section we provided them with feedback about what we had learned about them and where they sat on a visual scale. We asked them to self-validate their scores and to adjust their position if they did not feel it was right. (In the prototyping phase this really helped to check and refine the accuracy of each measure)

We were able to demonstrate that giving feedback as they went along, rather than at the very end (as some other methods use) had measurable impact.

We conducted a test and control experiment with and without these feedback elements and found that its inclusion increased answer thinking times by 40%.

We have now started to adopt this approach in survey feedback across a range of other surveys we run.
We can use this new instrument in a discovery stage of the client work, but also in segmentation, persona-building, journey map creation and creative or message testing by different personality, worldview, and the basket of cognitive styles tests.

For this project we conducted a personality profile of a group of 500 smokers and compared them to a control group of 500 non-smokers.

The smokers we surveyed proved to have quite a distinct personality profile in many dimensions.

The test predicted smoking with a correlation of 0.48. Basically half of smoking behavior could be explained by someone’s personality.
Key personality characteristics exhibited by smokers were:

- an external locus of control, (they were more fatalistic)
- lower levels of emotional stability
- more easily experiencing negative emotions and so actively avoided emotional engagement
- very promotion focus orientated mindset i.e. more rebellious, less likely to stick to the rules.

It is crucial to match creative executions with personality and thinking-styles. Think of it as a match for a decision-making genome. Conventional wisdom for anti-smoking messaging is to scare smokers with scary or even disgusting graphic images. They are emotionally-charged visuals highlighting the dangers of smoking. But if you decode smokers (“The Hidden Who”), you start to understand why so many graphic messages are not effective. When smokers test to be avoidant for negative emotions, they reject or skip over negative imagery. A number of behavioral studies have shown a backfire effect (“reactance”) among smokers to such scary, graphic warnings.

The way that smokers deal with these types of messages can be to blank them out. We tested this by showing a group of smokers and non-smokers a series of creative executions. We included graphic anti-smoking messages in amongst them, and timed how long they spent looking at each one. The smokers would click the “next” button much faster when shown one of these scary ‘dangers of smoking’ ads.

We then experimented with some alternative advertising messaging, specifically designed for the distinctive personality of smokers.
To appeal to their promotion-focus mindset, we designed some creative messages that re-positioned the process of giving up smoking as an act of rebellion and as a challenge.

Smokers rated these messages much more highly than the previous negative ads we tested.

Next we conducted some comparative dwell time testing. We found smokers spent longer time than non-smokers looking at these messages.

To non-smokers, this different style of ads had less meaning and so they clicked past them quite quickly, but they resonated with smokers and clearly encouraged them to think.

We have now rolled out this new “Hidden Who” methodology of decoding individuals at scale to decode why some parents hesitate to vaccinate their children, and what makes a superior performer among sales staffs.
C. GRAVES:

To put it all together…

Because marketing and communications research suffered at the hands of individuals who, as David Ogilvy said “don’t say what they think and don’t do what they say,” we have invented a new research and insights approach based on deep, behavioral science findings across more than 10,000 studies.

1) We have distilled these down to three lenses we deploy:
   - Personality traits
   - Cultural Cognition (Worldview)
   - Cognitive Styles

   These render a sharper, more effective picture of individuals at scale.

2) We have reduced traditional self-report bias with new methods of competitive choice and visualization

3) We have greatly enhanced traditional tools such as personas, journey maps and created the possibility for far more precisely-tailored content, creative execution, and message framing.

None of this replaces creativity. It better informs those creating the executions whether they be text, visuals, moving images, message framing, experience, or structural and choice interventions. Think of it as revealing a new decision-making genome against which we can map more effective, personalized content for better, healthier outcomes.
This page left intentionally blank.
THE REAL WHY & THE HIDDEN WHO
Fixing the weak links in how we measure personality to make better use of behavioral science in marketing. Deploying multiple behavioral science lenses to render a clearer picture of individuals at scale.

By Christopher Graves & Jon Puleston

ABSTRACT

Marketers, communicators and behavior-change strategists have lacked a powerful set of tools to boost effectiveness. They have depended on research, to be sure, but that research has been as blunt as a chisel when a proton beam is what is called for. Through a multi-year collaboration, Ogilvy & Kantar have created, tested and deployed a completely new instrument that more effectively decodes individuals at scale using a series of proven behavioral science “lenses” that had never before been pulled together into one comprehensive instrument. Together, these lenses decode personality traits, cultural cognition, and cognitive styles to reveal “The Real Why & The Hidden Who” aspects of individuals that truly serve as the drivers of and barriers to behaviour change. With the practical application of this instrument and its findings, we can now reinvent personas, segmentation, and inform a better, more empathetically resonant crafting of message framing and content. Just as advances have led to personalized medicine (treating individuals based on their own genome), this approach offers a novel, empathetic and much more effective way to move individuals at scale.

INTRODUCTION

Strategists, communicators and marketers depend on research to give them insights that can be used to effect a better outcome. Despite the thousands if not millions of hours spent on constructing award-winning ads we are still a long way off from developing advertising communication strategies that can drive real behavioral change at scale – see an end to obesity or convince everyone to stop smoking, encourage widespread engagement on climate change action, more complete uptake on vaccinations, secure retirement savings, engage employees. Often these are highly complex problems and what it takes to persuade one person will be different to what it takes change the behavior of another. Traditional approaches have either sought a killer campaign or segmented by demographics and stated preferences. Both are the equivalent of using a blunt chisel to do brain surgery when a photon beam is now available.

What holds back the effective delivery of so much advertising, communication and marketing strategies is an accurate read on what drives preferences, choices and behavior among sub-segments of audiences, consumers or patients. To understand this relies upon having effective ways of measuring and mapping out how individuals really think (as opposed to what they say), how their hidden personality traits, identities and worldviews filter everything they take in, and how their individual mindsets and cognitive styles nudge and guide how they make sense of the world.

As David Ogilvy once reportedly quipped (no doubt in frustration): “Consumers don’t think how they feel. They don’t say what they think and they don’t do what they say.”
We know from the work of neuroscientists such as Antonio Damasio ("Different Contributions of the Human Amygdala and Ventromedial Prefrontal Cortex to Decision-Making," Damasio et al, 1999) that most human decisions are strongly influenced by the emotional governors in the brain rather than reason. Damasio proclaimed: "We are not thinking machines. We are feeling machines that think."

We know from the Nobel Prize winning work of Daniel Kahneman and the behavioral economics field he co-founded, that human decision-making is guided by an array of cognitive biases and heuristics below our conscious level and making our species seem Predictably Irrational as the behavioral scientist Dan Ariely has written. The behavioral scientist professor and author Jonathan Haidt in his book The Righteous Mind summed up the relationship between emotion and reason in the brain as one of unequal partners; one calls the shots behind the scenes while the other trots out an explanation. “Reason is the press secretary to the emotions,” says Haidt.

All of which is to say David Ogilvy was essentially right and we cannot really trust what people tell us about their own reasons for preferences, choices and behavior. Using real behavior (such as “digital breadcrumbs”) can be helpful, but it looks backward and may be context specific. Meanwhile behavioral science points to the ineffectiveness of all logical and rational plans to change behavior.

THE BIG CHALLENGE: We needed to decode humans better at two levels: as a species; as individuals (at scale) leveraging behavioural science findings beyond the “nudge.”

SO WHAT’S A BETTER WAY TO DECODE HUMANS?

While the cognitive biases delineated by behavioral economists rule our whole species in general, our specific behaviors, individual choices and preferences are mediated by our individual makeup.

This paper tells the story of our attempts to develop a more robust multi-dimensional personality and cognitive decision-making style measurement tool to provide a clearer picture of human decision-making.

We wanted a test that could combine the 3 core, established behavioral science “lenses” used the view the human condition: personality, cultural cognition (outlook on the world) and cognitive (thinking) styles.

THREE LENSES TO REVEAL THEM ALL

Personality Trait Science (Big 5, Five Factor, NEO)
For nearly three decades, a wide and deep body of research has duplicated and confirmed personality trait factors as tightly correlated to preferences and behavior in many realms: from financial, to health & wellness, to environmental, risk propensity, to entrepreneurial success, sales success, management and leadership effectiveness, to consumer affinities (or dislikes) of shopping, to travel, and many more. An individual’s personality is roughly half hereditable and remains pretty fixed from young adulthood. Knowing how certain personality trait profiles tend to respond can help communicators and marketers better resonate with sub-segments (see “What Marketers Should Know About Personality-based Marketing” co-authored by one of us in Harvard Business Review).
Cultural Cognition

Initiated by the Cultural Cognition Project at Yale University led by Dan Kahan in the United States, Cultural Cognition evolves earlier theories such as the Cultural Theory (Grid-Group). It maps people on a grid and reveals their inherent worldview. From that worldview flow many ramifications flow.

Cognitive Styles

This is a basket of tests used individually by the scientific community. Each reveals proclivities and can predict preferences and behavior. They include:

- Regulatory Focus
- Promotion vs Prevention mindset
- Locus of Control
- External vs. Internal
- Time Perspective
- More past, present or future minded?
- Self-efficacy
- Do you believe you can succeed?
- Need for Affect
- Prefers emotion and visual narratives
- Need for Cognition
- Likes analysis, probing
- Hedonic v Utilitarian
- Consumer motivations

In a multi-year collaboration, Christopher Graves (Ogilvy Center for Behavioral Science) & Jon Puleston (Kantar Profiles) have created a comprehensive new research tool to do just that, using three newly-integrated lenses crafted from deep silos in behavioral science studies.

WHERE IT ALL STARTED: The problems encountered using classical personality measurement methods for commercial research

Two years ago, Ogilvy approached Kantar with the idea to try map out the personality and cognitive thinking styles of the Kantar panellists, to enable Ogilvy to perform more effective cognitive segmentation work. In light of the Cambridge Analytica scandal, they wanted to ensure that this work was undertaken within the clear ethical boundaries established by the market research industry.

Our initial approach was to take a range of well-established classical personality tests and field them to our panels. We did this on quite a large scale: for example, 20,000+ people from around the world were profiled with the standard OCEAN Big 5 personality test.

However, as we began to try to use these tests in real-life commercial research projects, we started to encounter challenges.

- The classical tests often struggled to differentiate one consumer from another: When we applied these personality tests to real-life projects, we often struggled to find any real differentiation between different personality segments. The example chart below, from a project to understand the personality of car buyers, illustrates the challenge. We were able to measure clear differences in only one of the Big 5
personality dimensions, while in the rest the data was completely flat. As a result, that data was not very useful.

- **The tests do not travel well**: What is more, when we ran these tests in different countries, the data proved very inconsistent and difficult to compare.

- **Single tests were unable to tell the whole story**: Each test in itself was not very useful considering the complexity of consumer decision-making. Just knowing the Big 5 facets of someone’s personality, for example, was often not enough, rather like going into an optician and they only have one pair of glasses with a fixed focal length. Every commercial project demanded that different facets of consumer personalities or decision-making processes be measured and compared.

We realised we needed to develop a more effective and comprehensive means of measuring personality and cognitive decision-making styles for commercial market research purposes and our journey started with having a good hard look at some of the existing methods used to measure personality, to find out why they often deliver poor audience segmentation.

**Part 1: Understanding the common challenges with classical personality measurement methods**

Most of the common means of measuring personality and cognitive thinking styles have been devised by academics in Western markets. They are isolated personality measurements that have often been tested only on small-scale audiences, in many cases on students in single countries, mostly in America.

So to help us understand how these tests were working we took a variety of them and fielded them at market research scale alongside each other, on a range of our panellists in different countries and at different ends of the cultural spectrum, to try to understand some of the common issues and understand how indivual tests were interrelated.
We tested out both the long and short-form version of the OCEAN Big 5 Personality Trait test alongside the Zimbardo Time Perspective and Boyd Past, Present, and Future focus tests and some classic tests of Locus of Control and Regulatory Focus.

We would not wish to denigrate the tremendous thinking and groundwork put into the development of these tests, but examining them purely from a commercial market researcher’s survey design perspective, and having fielded them internationally at scale and getting the same people to do the different test alongside each other, many of the issues were clear to see.

1. The challenges of self-validating assessment methods

Nearly all the personality tests we examined rely, one way or another, on fairly direct forms of self-assessment, essentially asking “are you like this?” which is a magnet for cognitive bias.

Analysis from across these experiments showed that self-assessment is fine for measuring what might be described as neutral aspects of our personality: for example, extroversion. People are able to identify themselves as being extrovert or introvert, and feel comfortable thinking of themselves as one or the other, but are less reliable at assessing aspects of personality that require self-criticism. The classical Big 5 test, for example, has a particular problem with self-assessment bias in measuring conscientiousness, openness and agreeability. As the chart below illustrates, few people are prepared to admit to themselves that they are unconscientious, closed-minded or disagreeable.

With such a large natural misbalance of answers, it becomes difficult to differentiate people in these personality dimensions. So it should be no surprise that it was in these aspects that we had most difficulty in differentiating consumers for commercial projects.

2. Overreliance on repetitive Likert scales

Compounding this, nearly every test we examined had an overabundance of Likert scales, often employing banks of twenty or even thirty at a time.

When in some doubt, a significant proportion of people will say they moderately agree about just about anything, and the problem only gets worse if respondents are not feeling engaged. When testing, we found upwards of 35% overlap of mutually exclusive answers to separate questions, such as the example below.
This problem became even greater in certain Asian countries, both because of a culture in which respondents tend to agree with things more, and because the abstract nature of the questions carried an increased risk of misinterpretation. In India, we found up to a 50% overlap in some cases of mutually exclusive questions.

This creates a lot of noise in the data, making it difficult to differentiate one personality characteristic from another.

That’s not to say Likert scales do not have a critical role in measuring personality – the challenge comes when they are overused and pose difficult questions that confuse participants, especially after translation into different languages. They have to be used carefully, with additional care taken to ensure participants give them their full attention.

3. Modal biases

We also observed a more practical challenge with some tests in rendering larger-range scales on mobile devices. The number of options meant they needed to be vertical ordered on mobiles, which delivered much higher top-scale bias (a 15%+ difference) compared with questions laid out left to right on larger-screen devices.

This can exaggerate some age-based personality biases. For example, the increased positivity of young people compared to older people can be exaggerated by the device the survey is completed on, with more young people using mobiles, and more older people PCs.

4. Western-centric questions

Another major issue was the “western-white-male-wealth” biased nature of the questions being asked in some of the tests we evaluated, such as the Regulatory control survey. This test is used to assess attitude towards taking risks, and included question about…
- Going camping in the wilderness
- Investing 10% of your annual income in a moderate growth-diversified fund
- Betting a day’s income at a high-stakes poker game
- Going down a ski run that is beyond your ability
- Going white-water rafting at high water in the spring
- Taking a skydiving class

All of which would be outside the range of experience or consideration of most people living outside the collegiate world of the US.

5. **Personality metrics that are one step removed from being commercially useful**

The academic nature of many of these tests, never originally designed for any commercial purpose, means that some of the questions and personality constructs are not very transferable to understanding real-world consumer decision-making processes.

Take, for example, “Emotional stability”, one of the OCEAN Big 5 measures. This reflects how strongly one feels and acts upon emotions, but also involves a measurement of someone’s latent anxiety levels.

A highly emotional decision-maker might make quick, impulsive choices, but an anxious decision-maker might want to invest time and thought to ensure they make the correct decision. So one generic measure of emotional stability is difficult to use – there is a need to subsegment this personality type if it is to be any use for marketing purposes.

Attitude towards risk is another good example. These surveys focus on physical/ action-based risk, but a consumer marketer is more interested in consumer decision-making risk, for example how much information someone needs before being prepared to make a decision. Knowing whether someone likes white-water rafting is little help in answering this.

6. **Uncontextualized generic questions**

This leads to the observation that many questions in these personality tests are very generic in their nature, and not anchored to consumer-based decision-making processes. To be asked, for example, if you feel you are a “conscientious person” is somewhat meaningless without some sort of context, resulting in less useful responses. In real life, we exhibit different levels of conscientiousness depending on circumstances. I am likely to put more thought into buying a car, for example, than washing-up liquid.

7. **Standalone tests which only measure isolated facets of someone’s personality**

Any one test only measures one facet of someone’s personality, but all the different tests stand isolated from each other, when tested together we were able to see clear cross correlations and relationships between different tests, down to the interwoven nature of different aspects of our personality but in their existing form there is no way of linking these tests together efficiently.
Part 2: Undertaking a ground-up rethink about how to conduct personality measurement

Armed with these insights, we started to explore how to address some of these issues by experimenting with different approaches.

2.1 Tackling confirmation bias by switching from using Likert scales to Competitive Choice-based Prioritisation (CCP): Our first move was to think about ways to reduce reliance on Likert measurement techniques by using a competitive choice approach. This is a technique where instead of asking people to monadically evaluate themselves against a series of personality characteristics, we present a cluster of competing characteristics and ask people to simply pick out the ones they feel most apply to them.

We originally pioneered this approach to more reliably measure issues that consumers naturally over-claim about, for example consumer attitudes to sustainability-related issues. Most people, when asked to assess the importance of the government doing more to tackle global warming, would say it’s very important. If they are instead asked which issues are important for the government, and are given a competitive list which includes global warming among others like gender equality, improving the health service etc, without constraining the number selected we found the proportion who choose global warming decreases to more realistic level and reveals those who really think it important.

Likewise, when you ask people to pick out their personality traits from a competitive set, they pick the ones that they can truly identify with, and we found this significantly reduced the overlap between mutually exclusive personality characteristics, greatly reducing the amount of noise in the data.

Furthermore, because it is quicker to select options like this, we found that we could evaluate nearly three times as many personality dimensions in the same timeframe as when
using Likert scales. This is important, as one of our goals was to combine multiple tests into one measurement instrument, meaning we needed to ask more questions.

2.2 Dealing with underclaim: This approach did not completely solve the problem, however. While it reduced overclaim, with less people claiming to be self-disciplined, open, calm and sympathetic, it did not tackle underclaim. Few people are prepared to tick a box that identifies them as disorganised or critical, so these measures did not change.

To assess these more “negative” aspects of our personality, we needed to use different approaches. Our experiments uncovered several very effective techniques.

2.3 The Silent dog method*: This method, named and championed by Ray Poynter, comes from the idea of examining what is not said. We realised that if we transposed some of the questions into the opposite dimension, and examined those who didn’t select them, we could get closer to the truth.

*I am untidy (12% tick)

*Sherlock Holmes solved a murder mystery by noticing that the dog did not bark so must have known the murderer.

2.4 Family/friend anchoring: While few people are prepared to concede that they are critical or quarrelsome in a general sense, we found they were a lot more able to evaluate themselves as such in relation to people they knew. “I am prepared to admit that I am moodier than my brother.” A subtle change of wording encourages comparison not to some abstract average, but to family and close friends, and this goes some way to improve the levels of negative self-reporting.
2.5 Adding some retrospective perspective: In similar vein, we found that people were more able to look with less bias at their younger selves than their present-day selves.

“I won’t admit to being the type of person who is late to meetings, but I am prepared to concede that I was often late for school.”

As so many aspects of our personality are formed in our youth, this approach provided useful and revealing insights.

2.6 Using memes to encourage more honest self-examination: We found we could improve negative self-reporting still further by highlighting the difficulty people have at observing some aspects of their personality.

We communicated this using the famous Grumpy Cat meme and a narrative explaining how hard it is to self-observe. So successful was this technique, the grumpy cat R.I.P. might well be due some sort of special retrospective market research honour, since nearly twice as many people self-reported as being moody, and every negative dimension we measured increased.

2.7 Using behavioural measures*: The next technique we integrated into our methodology was a switch to using more behavioural-based measures that are a little easier for people to answer truthfully. For example, only 12% of us are prepared to concede to being disorganised, but 50% are prepared to admit that the clothes in their bedroom drawers are disorganised.

*The value of behavioral question techniques have been outlined in a previous ESOMAR paper: The Segmentation Revival by Puleston, Brownlee & Wheatley 2018
We cannot assume that just because someone’s drawers are untidy, they are a disorganised person and thus have lower levels of conscientiousness-- all it provides is a small clue. We found that by asking several such questions they can all add up to a powerful means of measuring personality aspects we find harder to confront directly.

We also realized that the more we could make these behavioural-based measures situationally relevant to consumer decision-making the more useful they would be for commercial research.

And we needed to indentify behavioral measures that would allow us to sub-segment each of the personality measures to ensure the questions covered all the different personality aspects relevant to real-world decision making.
To devise our re-imagined test, we evaluated literally hundreds of different types of behavioural measures, in a series of experiments like the examples below.

At home: Everything in my clothes draws is neatly organized
At home: I get annoyed by mess
At home: I like to tidy up
At home: I made the bed this morning
At home: I put my music or books collection in alphabetic order
At home: I sometime leave my clothes on the floor
At home: My bedroom right now is very tidy
At home: My partner sometimes moans at me for being messy
At home: There are likely to be things in my fridge past their sell by date
Cooking: I always try to follow the recipe
Cooking: I put a lot of effort into cooking when I do it
Cooking: I wash up straight after every meal
Cooking: I watch how much I eat
Cooking: We plan out meals in advance
Habit breaking skills
I have successfully lost weight
I always try to read reviews before choosing what film to watch
I am a bit of a gambler
I am very careful about recycling
I collect things
I double knot my laces
Money: I am organized with money
Money: I have a monthly budget that I try to stick to
Money: I regularly check the transactions in my bank account
On holiday: I unpack my suitcase and put all the clothes away
School: I always arrived on time for school
School: I always handed in my homework on time
School: I occasionally took a day off pretending not to be well
School: I tended to do my homework at the last minute
School: I was a bit lazy when it came to preparing for exams
Shopping: I always cut out and use offers when I see them
Shopping: I always make a shopping list before I go to the supermarket
Shopping: I am an avid reader of reviews to help me decide what to buy
Shopping: I like to keep receipts
What is your attitudes towards voting?: I always vote in every election
What is your attitudes towards voting?: I don’t always vote
What is your attitudes towards voting?: I never really vote
What is your attitudes towards voting?: I voted in the last election
Work: I am exacting in my work
Work: I have never called in sick as an excuse to take a day off work
Work: I like to follow a schedule
Work: I like to get into to work early

We then used principal component analysis to identify the most stable and reliable set of behavioural measures assess each personality dimension.

E.g. Conscientiousness

<table>
<thead>
<tr>
<th>Low Conscientiousness</th>
<th>High Conscientiousness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regularly vote</td>
<td>79%</td>
</tr>
<tr>
<td>Written a will</td>
<td>67%</td>
</tr>
<tr>
<td>Double knot my laces</td>
<td>42%</td>
</tr>
<tr>
<td>Stick to monthly budget</td>
<td>30%</td>
</tr>
<tr>
<td>Always make a shopping list</td>
<td>26%</td>
</tr>
<tr>
<td>Keep shopping receipts</td>
<td>26%</td>
</tr>
<tr>
<td>Looking at ingredients on label</td>
<td>16%</td>
</tr>
<tr>
<td>Put a lot of research into buying car</td>
<td>16%</td>
</tr>
<tr>
<td>Saving money for the future</td>
<td>30%</td>
</tr>
<tr>
<td>Fold headphones before putting in pocket</td>
<td>30%</td>
</tr>
<tr>
<td>Make bed in the morning</td>
<td>40%</td>
</tr>
</tbody>
</table>

Pulling all these techniques together, calibrating and aggregating each element to create we were able to create a much more balanced and rounded picture of each personality trait that allowed us to differentiate real-life behaviours far more effectively than we could achieve with the classical personality measurement techniques we had evaluated.
Part 3: Combining multiple personality & cognitive tests into one survey

The final part of the story was working out how to efficiently combine the core personality, cultural cognition and cognitive thinking styles we wanted to assess into one survey, to get a full 360-degree view on someone’s personality and anchor these to their underlying consumer decision-making style. We identified 8 specific tests we wanted to combine into one survey.

3.1 How we integrated the tests: Through a series of pilot experiments, we combined collections of these tests we had developed and adapted using the techniques explained above, then forensically examined the contribution that each question from each test made to these combinations, and the underlying correlation between questions from the different tests.

What was clear from these experiments was that the different personality aspects are often closely interwoven. For example, a person who considers the future and has a strong sense of internal self-control is also likely to be more conscientious, so the answers from each test could be used to cross-validate each other, providing greater data stability overall.

By doing this, we carefully whittled down each test to its core unique elements, removing overlapping questions and using the answers from one test to inform the answers to another.

3.2 Blockchain validation: Working a little like a blockchain validation process, we found that by combining many personality tests into one survey meant we could use the answers from every part to independently cross-validate and calibrate the answers to every other part of the survey. This made us less reliant on one set of questions from one test to predict any one personality dimension, and meant we could shorten each part of the survey, since the other parts of the survey could fill in the gaps.

In this way, we were able to combine these eight independent tests, containing 30 minutes of questions, into one integrated survey that only took around 17 minutes to complete.
3.3 **Designing the survey experience**: Recognising that 17 minutes was still a long time to ask people to concentrate, our main focus from a survey design point of view was to work out how to most effectively hold respondents’ attention throughout the survey. To prevent repetition fatigue, we mixed and matched different question formats, with no more than 8 questions in any one repetitive loop. To reduce modal effects, we carefully designed the range-type questions we used to work consistently across devices.

We designed the survey around a modular format, breaking up the component parts of the survey, into three-minute-long, “thinking chunks”. We began each with what we describe as a “thought starter”, a question to introduce the topic of the next section and grab the respondent’s attention.

At the end of each section, we gave the respondents feedback about what we had learned about them along the way, and asked them to validate the accuracy of our assessment. If they thought it was wrong, they were given the option to correct it. This feedback really helped in final-stage piloting to ensure the survey was functioning effectively in the minds of the people taking it.
We were able to clearly quantify the value of these engagement elements in a research-on-research experiment, where we removed them to compare completion rates. The thought-starters and feedback reduced dropout, significantly improve the time respondents spent thinking about their answers, and increased willingness to continue participating at the end of the survey.

3.4 Piloting to ensure our personality test worked across cultures: We finally undertook two waves of large-scale multi-country piloting, to refine the survey to ensure that the results it delivered were consistent across countries. This involved quite a few detailed refinements to some of the questions. We found it almost impossible to ensure that every question worked in every country, however we found that one of the other big advantages of basing each personality measure on answers from across multiple questions from different parts of the tests (up to 40 different measures contribute to each personality dimension) was that it provided much more stable cross-country comparisons than any single test method could offer. If one particular question was subject to some cultural variation it would not completely corrupt the whole score. With the final version of the test, we were able achieving correlations of c0.9+ between countries in the majority of personality dimensions.

Part 4: Putting the survey tool to use

In the final part of this paper, we would like to show 5 case studies to illustrate the range of way we have been able to start using this new approach to personality measurement in practice to get a clearer understanding of audiences and as a result devise more effective communication strategies.
Case study 1: Measuring the appeal of different styles of advertising to different personality groups

To test out the personality test, one of the first projects we undertook was to ask people who had completed the test to then evaluate a range of different ads. We found we were able to quite clearly differentiate their appeal to different personality groups.

We could also see how pure advertising design could shape the appeal of an ad to different people with different thinking styles. In this example, two identical messages were presented in different ways. One appealed at a more emotional and hedonic level than the other.

We could also differentiate the impact of the core messaging. In the example of testing two finance ads, the one that focused on taking care of money appealed to people with a personality mindset that emphasised an external locus of control and prevention. The other, which emphasised growing money, appealed to a promotion-focused and more open personality group.
Case study 2: Using personality evaluation to help devise more effective anti-smoking messages

To demonstrate how we could use an understanding of the personality of a target audience to develop more effective advertising messages, we undertook a project for a client wanting to learn more about smoking cessation. We mapped out the personality of smokers and those desiring to give up, and discovered they had a clear personality footprint. The personality test was able to predict if someone was a smoker with $0.48$ correlation.

This personality footprint exhibited lower levels of emotional stability, more fatalistic viewpoints, and very promotion-oriented attitudes. Smokers had a strong tendency to avoid negative emotions. Interestingly, the typical anti-smoking message focuses on provoking strong negative emotions, so an understanding of this personality type shows why so many smokers simply block out these messages. Our experiment showed that smokers actively avoided looking at this type of ad, clicking past them much faster than non-smokers.

These insights led our team to devise some different styles of messages, more suited to the personality of smokers – more solution based. These focused on external locus reasons for giving up, and positioned the idea of giving up as a challenge, to appeal to a smoker’s promotional mindset.

Testing these new types of messages, we found smokers to be far more likely to be engaged in reading them. They spent up to $50\%$ longer before clicking the next button, compared to the typical smoking ads, and more time than non-smokers looking at them.
Case study 3: Using personality testing to understand the barriers to mothers getting their children adequately vaccinated

This project was undertaken for a client in Asia to understand more about mothers and their attitudes towards vaccinations. The core brief was to understand the barriers to mothers getting their children adequately vaccinated, and what drives the willingness of some mothers to fully vaccinate their children beyond the local government minimum compliance. A customised version of the personality survey was fielded in five countries across Asia, integrated with a bespoke range of questions around parents’ attitudes towards vaccination and their child’s wellbeing.

From the results, we could clearly identify the core personality traits that correlated with a confidence about vaccination, versus those that were hesitant about getting their child vaccinated.

**Building personas:** From this, Ogilvy was able to carefully devise two cognitive personas to help understand these different personality groups, which were used in workshops across Asia to devise strategies to deal with the challenge of communicating to them.

**Making cross-cultural comparisons:** We were able to map out and contextualise cultural differences between countries and identify unique personality differences between countries that impacted on attitudes towards vaccinations. One example is the differences in the levels of optimism bias exhibited by mothers in each country, which we discovered significantly affected whether they were worried about the dangers of not getting their child vaccinated. We also observed different levels of social influence regarding the topic, and how doctors and health-care professionals engaged with mothers in each country.
Case study 4: Undertaking a cognitive segmentation to understand the barriers to losing weight

For a client marketing a severe weight loss product, we embarked on a deeper understanding of those struggling to lose weight.

Doing an 3-lens analysis of personality traits, outlook, and thinking styles of those who struggle most with weight management uncovered a correlation of personality traits (low conscientiousness & low emotional stability) coupled with a mindset of external locus of control and low self-efficacy-- a lack of belief in solutions.

The creative strategy that evolved from understanding the personality profile of people struggling with weight management was a “New Beginnings” test creative execution that championed easy-to-do, small victories, leveraging their external locus of control.

Recognizing that people will not attempt weight loss unless they could believe they have the competence and self-confidence (self-efficacy) to succeed, the campaign emphasized building self-efficacy step by step.

The proposed campaign outperformed side-by-side alternatives in every market tested in ad testing.
Case study 5: Using personality testing to understand the personality of a company to its customers

A client that depends on a large, independent sales population wanted to better understand (beyond demographics) who these sales people are and what motivates them—particularly the high performers—so the company could recruit and retain more like them.

The survey was voluntarily and anonymously completed by the sales people across the organisation, and it was also sent to a cross-section of their customers as well.

The results provided deep and broad ranging insights. It highlighted the strongly extraverted and promotion-focused nature of the sales organisation.

It helped them to more clearly understand what it took to succeed in their business, and revealed how the workloads of the middle and senior tier of management made it difficult for them to plan and see beyond near-field deadlines and the instinctive decision-making nature of the senior management team.

The study also revealed some fascinating personality differences between their customers and the tiers or management, that frankly we are still in the process of making sense of…
SUMMARY

Because a robust and growing body of behavioral science research has revealed that human decision-making and behavior cannot be explained by, or changed by, logical, rational approaches, we require a new approach to decoding individuals at scale.

Having a multidimensional (personality, cultural cognition and cognitive styles) measurement instrument like this new innovation can be compared to an optician having a large set of different lenses: It allows you to examine almost any issue with more clarity through adding relevant lenses. Each lens helps reveal a bit of the “Real Why and the Hidden Who” of individual behaviour, but used together they render the sharpest image of all.

For strategists, communicators, and marketers – indeed anyone looking to better understand the real drivers of, and barriers to, behavior change-- this presents a way to bring to life and practically apply behavioral science theory. It allows for a new approach to 1) diagnosis of the challenge; 2) segmentation of groups from their inside-out; 3) reinventing personas; 4) reframing messaging and recrafting content so that it resonates more effectively when matched to individuals at scale.

This paper describes a two-year journey to develop just one 17 minute comprehensive survey measurement instrument, but we have learned so much along the way. We believe that many of the techniques we devised to tackle the challenges of personality, cultural cognition and cognitive styles testing have applications in the wider world of research.

- Competitive choice-based prioritisation
- Honesty priming
- Silent dog techniques
- Behavioral-based questioning methodologies
- Blockchain-style cross-validation techniques to optimised surveys and provide more cultural cross stability
- Modulization of surveys and giving feedback throughout a survey to improve engagement.
- Moving marketing & communications away from blunt, outmoded approaches toward an empathetic resonance with individuals at scale

We know this unified and multi-disciplinary approach is just the beginning in reinventing many aspects of marketing and communications research that will result in much more effective segmentation as well as better-tailored messaging and content.
References:


Youyoua W, Kosinski M & Stillwella D (2015): Computer-based personality judgments are more accurate than those made by humans, Proceedings of the National Academy of Sciences


Appel M, Richter T & R Kepler (2010), Narrative persuasion and need for affect, the journal Media Psychology

Puleston J, Sleep D (2011): The Game experiments, ESOMAR congress


Find out about the personality & cognitive thinking style measures referenced in this paper:

OCEAN BIG 5 Personality: https://en.wikipedia.org/wiki/Big_Five_personality_traits
Cultural cognition: https://en.wikipedia.org/wiki/Cultural_cognition
Regulatory Focus theory: https://en.wikipedia.org/wiki/Regulatory_focus_theory
Locus of Control: https://en.wikipedia.org/wiki/Locus_of_control
Self-efficacy: https://en.wikipedia.org/wiki/Self-efficacy
Need for Affect: https://en.wikipedia.org/wiki/Affect_(psychology)
Need for Cognition: https://en.wikipedia.org/wiki/Need_for_cognition