ESOMAR | 70
WORLD RESEARCH

Asia Pacific 2017

Conference Papers

COPYRIGHT

Published by ESOMAR, Amsterdam,
The Netherlands
Edited by: Deborah S. Fellows

ESOMAR Office
Hoogoorddreef 5
Atlas Arena, Azië Building – 5th Floor
1101 BA Amsterdam
The Netherlands
Phone: +31 20 664 21 41
Fax: +31 20 589 7885
Email: customerservice@esomar.org
Website: www.esomar.org
ESOMAR Publication Series Volume S379 APAC 2017
ISBN 92-831-0294-0

Table of Contents

Table of Contents

Programme Committee

Lifeng Liu (Committee Chair)
Ipsos, China

Andy Zhao
GfK, China

Eri Yamazaki
The Coca-Cola Company, Japan

Grant Bertoli
Buzzebees, Thailand

John Smurthwaite
TNS, Malaysia

Ritanbara Mundrey
Nestlé, India

Sally Wu
BBC (Global News), Singapore

Editorial

Discoveries

Welcome to China, my homeland. Please allow me to tell you about China. On one hand China is a leading emerging country, particularly in the new economy, and on the other hand it is an information island. In many ways China is still disconnected from the world, as for example, Facebook is not here.

China is, like many developed and fashionable APAC countries or emerging APAC markets, a great market, allowing us to discover new generations, new consumers, new clients, new media, new retail, new marketing, new advertising, and new research, as well as new big data. Discoveries available, right here.

Due to the development of the internet, of course including mobile networks or WIFI using smart phone, tablets, etc., many theories and approaches collapsed as the foundation on which they were built was based on the previous environment. The new digital world needs us as researchers to discover.

How is retail moving to e-commerce?

In the past, consumers went to shops for shopping, the shops received products from wholesalers, and wholesalers were supplied by manufacturers.

Today, consumers go to websites for reviews, ratings, and price comparison. They then buy their chosen product through e-commerce and wait at home for the product to be delivered. They also contribute reviews and ratings during the entire purchasing process. How can manufacturers impact consumer decisions, how can traditional marketing approaches and research tools be used?

How are manufacturers changing? In part, manufacturers rely on research companies to understand consumers and then decide what products to launch or offer to the market.

Today or tomorrow, manufacturers might be designing personalized or tailor made products. They might rely on internet based product design applications to develop products. Consumers could become product designers. Manufacturers might need to depend on big data companies to reach target consumers directly. The era of big scale advertising campaigns and big scale production might end.

How is the advertising industry evolving?

In the past, manufacturers relied on ad agencies to develop and launch commercials to influence consumer decisions.

Today, manufacturers request that commercials be more accurately launched towards their targeted audiences, relying on big data and data mining platforms to decide where, to which consumer, when, and what key words to launch. Consumers have left their footprints and behavior information on the internet. The traditional advertising theories collapsed. The advertising industry needs to be re-built based on the big data environment. For example, when you are drinking alcohol, your wife or your girlfriend might have at the same time received a message reminding her to book a substitute driver for you through an e-substitute driving app. Scary?

How is the market research industry transforming?

In the past, we conducted research for clients to help them identify whitespace and opportunities, help them test concepts, products, and commercial copies, and track their brands, campaigns, and product performance. We also have our secret weapon, of which we are quite proud – we interview people to collect information by selecting respondents through random sampling.

Today, consumers express themselves through social media. We do not need interviews in every situation. Our clients or new technical companies can track information, including attitudes and behaviors. Random sampling is not necessary as the cost of universal data is cheap.

Consumers today are linked with manufacturers. Many manufacturers talk to their consumers directly, and target consumers can even help manufacturers co-create concepts, products, and brands. Why not?

I do not mean to say that research theories are also collapsing, but clearly they are fading. The good thing is at least we have the knowhow, we know how to analyze data, and how to develop insights from data to help clients make decisions.

Obviously, there is much waiting for us to discover:

Consumers are different now. They are not just consumers, but are also designers, they are “we media” owners, they are experts in judging product performance and services. One consumer, and even a group of consumers, might be wrong but consumers as a universe cannot be wrong. Internet makes it possible.

Our clients might have become lost. Many international FMCG clients are suffering in China, as they suddenly realize the way they market products does not work any longer. Consumers are smarter and more hypercritical. The competition is more challenging. They need us, or somebody else if we cannot, to help them find the way out.

Media is different now. Online video websites have more impact than TV channels in China, and we do not even need to mention social or “we” media. The process of developing commercials, identifying what commercial to develop and test, and even how to test are very different now. Do we even still need tests?

How products are launched could be very different too, as products could be co-created by the client and its target consumers.

There are new economy questions waiting for discovery. China by far is the biggest country in history moving from underdeveloped to emerging and to developed, moving from low income to high income. China has a huge middle-class population from emerging markets waiting for us to discover.

They are many new clients including BAT (Baidu, Alibaba, Tencent) and Didi (China Uber) waiting for us, and there is no need to even mention Facebook, Google, Airbnb, etc. Their marketing outlooks are very different, and they are waiting for us to discover this.

Enough! That is why we are here in Shanghai. We will experience the journey of discovery as we hear from clients, great talents from advertising companies, from media and consulting companies, and of course from technology companies.

I hope that by reading this collection of papers you too will enjoy discovering the new insights and information about the region.

Lifeng Liu
Programme Committee Chair

PART 1
Discoveries at the Heart of Client Transformation and Impact

Building Effective Networks

Leveraging the power of the ecosystem to deliver greater impact to business

Madhumita Chakraborty

Introduction

In the last five years, the Insights function in many organizations is undergoing significant changes. More and more organizations are adopting matrix structures, expanding their global footprint, merging or producing offshoots and grappling with emerging channels of reaching consumers. As a result the demands of the function are changing - both increasing in terms of volume of demand and evolving in expectations. Change is constant. But over the last few years the magnitude of the change, along with the tectonic shifts in organization structures, requires the Insights function to make some fundamental shifts in order to remain relevant to the wider audience.

As the Insights function undergoes changes, the impact percolates to all associated teams - most notably the research agencies that work with the Insights teams. To research agencies this presents both a challenge and an opportunity - the opportunity being one of widening scope of work, wider stakeholder exposure and access to C-suite execs. The challenges are equally great - keeping abreast of changing structural dynamics, upskilling talent that can speak the language of business impact and not just marketing, and in many cases having to compete for share of spend with non-traditional competition, i.e., with other types of agencies outside of the classic “market research” fraternity.

Through this paper I would like to share some learnings from PepsiCo India and other companies as these organizations navigate through these changes. By no means is this exhaustive or prescriptive, the thoughts in this paper are meant to spark discussion rather than provide pre-fabricated solutions.

Context

We live in a VUCA world (Volatile Uncertain Complex Ambiguous); but never more so than today. The average life of a Fortune 500 company today is 12 years, down from some 50+ years sometime back. The rapid pace of change and this need to stay relevant in consumer’s minds has led organizations to proactively rethink structures and ways of working within organizations. This cascades down to the Insights function in multiple ways. While there are different contexts and background for each organization, they can broadly be captured under three top themes for Insights function:

1. The widening scope of the Insights function
2. Drive for actionability
3. Drive for productivity

The widening scope of the insights function

Executives in many large corporations are realizing the potential of alleviating the Insights function into an independent organization, reporting directly to the CEO or COO’s office. A longstanding dream of most Insights teams is to be recognized as an independent organization that can drive wider business decisions without beingovershadowed by the marketing function that has been its mainstay of existence. This transition brings with it a new set of stakeholders - from Commercial, Human Resources and Production - all of whom are eager to tap into the insights tools and resources; but who at present have very different perspectives and very little appreciation of research methodologies. Additionally, many large corporations are also trying to balance the Global- Local connect for sharing and optimization. While the advantages are super clear, this results in organizations and Insights teams having to resort to new ways of working, calibrating methodologies and revisiting agency partnerships.

Drive for actionability

Insights teams across the globe are being tasked with the charge of improving their speed of response, becoming more predictive in their workstreams, finding ways to seamlessly merge multi source data and moving to standardized tools across markets. This is not unique to organizations undergoing transformation; but is definitely heightened in organizations undergoing change. This is complicated because the "call to action" now comes from multiple stakeholders with whom the classic insights team was not accustomed interacting.

Drive for productivity

Again not an unique ask of Insights, but whereas previously corporations focused on productivity much more from media and marketing budgets; with insights called out as a separate entity within the corporation, the ask for productivity now also extends to insights function. This manifests itself in two forms - greater leverage of existing assets and data tools and repurposing budgets from reporting back related activities to forward looking or mix development activities.

In summary, the changing forces within large corporations are tasking the Insights function to Do more, Be Different and Imbibe New ways of Working. The upside of this is definitely very positive, the challenge lies in building bridges quickly to gain appreciation of the wider business issues and in leveraging the power of existing networks to deliver to these new and evolving demands.

The journey of a thousand miles begins with one step (Lao Tzu)

The journey that PepsiCo undertook, and indeed many companies have taken similar steps, is by no means easy. It took time, patience and above all a conviction that we were on the right path. Sometimes things unraveled and had to be reconstituted, and there were times when the vision was clear but to fructify the vision we had to seek support from unaccounted quarters.

The following reviews some of the significant steps undertaken. While not necessarily in order, there is definitely a clear evolution as we went along over the last two years.

The first significant measure was to map WHO we now catered to**.** *Defining the stakeholders* within the organization was the starting point in organizing our teams and priorities. In the first year the stakeholders would still largely be the Marketing team and Senior Management. But over the next year or two we would consciously take measures to expand our sphere of influence with the Sales and Finance teams; and subsequently with Human Resources and Production. Again there is no right or wrong way to this, but helping prioritize and building a roadmap enabled the team to stay focused on delivering high quality deliverables and building a reputation for results in the wider organization.

To bring this to life, we embraced a *Matrix structure* within the Insights function to replicate the wider organization; calling out CMI teams that serviced respective Business Units; and a functional service line in Analytics and Shopper that spawned support across all BUs. In principle this sounds logical and easy. As we rolled out this structure we realized the importance of clearly defining roles and communicating the same across the organization to enable clarity on who does what, first within the Insights teams and the wider corporation. This is easier said than done. To start with this took the external world, particularly the research agencies interacting with the teams, some getting used to. It added the number of stakeholders they had to interact with, and the buy in process for landing decisions was longer in the first few months. As the structure stabilized, and through repeated streams of communication on people's roles and responsibilities, we were able to ultimately smooth this out. A clear enabler to this was to keep the agency partners abreast of the transformation journey.

The second step in this journey was to establish the *Ways of working*. In the early days, even as roles were defined, there was still some handholding required to nail out the processes and response mechanisms. While deliverables were always met, there was at times double communication, and at times the reverse to land the response mechanisms. We worked with the HR teams to bring in some processes and roll out classic frameworks like RACI (Responsible- Accountable- Consultant- Inform) and DICE (Decision Maker- Inform- Consult- Execute). This helped bring clarity on how people collaborate and respond; helping the insights team improve its overall speed of responsiveness.

With the basics in place, about six months to a year down the line, we began the concerted effort to *widen the sphere of influence* to other stakeholders in the organization. As a starting point, the insights team started attending cross functional meetings – even when it did not have an active role to play - as a measure of induction and making ourselves aware of the perspectives of different stakeholders.

Next, within the team we allocated Insights FPRs (First Person Responsible) for all key stakeholders - including the Commercial Unit heads - and began interacting with them to see if we could extend Insights support to their key planning cycles. To cement this, the Insights FPRs started sharing monthly performance reports and interesting reads. In the first few months there was little appreciation for the same. But over time, the *Commercial Units* starting recognizing the value of this support, and gradually worked with the Insights FPRs to build mechanisms that have helped the Commercial units build consumer centricity and made Insights a core partner in their decision making processes.

An enabler to this was a series of interactions and training programs. *"Insights for Non Insights*" was a series of customized interventions for the Commercial Units to build appreciation of Insights tools and skills in Data Triangulation. Not only did this series make the frontline more connected and aware of Insights processes and tools, it has helped the teams improve their ability to make data backed hypotheses and discussions. Another measure undertaken was to build a house for all Market Walks and mine this for emerging trends; thereby establishing a knowledge management mechanism that enabled frontline staff to get actionable consumer, category and competitive intelligence for their markets.

With the desire to impact the wider organization, outside of the Sales system which was the first port of call, we next honed in upon two other critical partnerships – the *Information Systems team and the R&D team*. The Information Systems team is the custodian of large database of Customer Sales and Experience data. While the database is expansive, it can be quite daunting to navigate through the mass for actionable information. Over the course of six months, the Insights and Information Systems teams have worked closely to redesign reporting interfaces and data navigation rules to enable wider socialization of internal data, thereby building common language within various functions in the organization; and reducing dependence on external data sources/ primary research to support decision making. By itself, this is a standard protocol in most organizations, what makes it unique is that this partnership remains ongoing and is now extended to special projects; which has helped bring out the best of both the IT and insights colleagues.

Insights has always had a strong rapport with the R&D team, as Innovation development and testing is a critical research area in most CPG companies. As a step up to this, Insights is now partnering with R&D on updating normative databases, relooking at integration of Sensory - Consumer and working on preparing learning sessions from Innovation journey. This looks simplistic, but building this cohesiveness requires teams to revisit and align their protocols, both internally but also with external partners and across global teams.

Change management is never easy. To enable Insights function to remain forward looking and to keep the internal dialogue going, *Insights Task Forces* were set up. Some focused on functional tasks and some focused on softer skills and ways of working. The continuous dialogue and conscious call outs have helped people during the transition process until this becomes the new norm and ingrained in people's ways of working. Alongside this, there has also been significant investment in upskilling Insights talent on core functional and leadership skills. On an average people have spent at least five days per year on leadership training and 5-15 days on functional training programs.

Coming together is a beginning, keeping together is progress and working together is success (Henry Ford)

Enabling change would not have been possible without the partnership of the research agencies that partner Insights teams. This section focuses on some of the steps undertaken by PepsiCo to leverage the partner network on this journey.

One of the most applauded processes initiated by PepsiCo India is its *Annual Agency Review and Partner meet*. This meeting series has now been running for five years, and has evolved over the years. In the first year it started out as an annual meeting between Agency Senior Leaders and the Insights team to objectively discuss the quality of the relationship over the year, enable a forum for two way dialogue, and to provide some visibility into plans for next year. Over the next few years this evolved and has become a landmark event for agency partners and PepsiCo Insights alike.

Some of the salient features of this meet are

- Annual feedback sessions between Agency Senior management and PepsiCo Insights leaders
- Two way dialogue, conducted in an offsite/ neutral location
- Feedback compiled in a robust manner – includes feedback from insights team, end users and calibrated internally before sharing with agency teams
- Review scores accompanied by action plan shared by other team (I.e., PepsiCo Insights shares action plan on how to meet concern areas highlighted in Agency feedback to PepsiCo and vice versa)
- Follow up quarterly governance calls on tracking against action plans
- Exposure to senior PepsiCo leaders on their expectations from partners
- Share out session- with PepsiCo Insights leadership sharing out Annual Plans with the leadership of Agencies, thereby providing visibility into plans and inviting thoughts on how to partner
- Agency Reward and Recognition night- a rewards function with a rolling trophy for Best Agency and other awards for teams and projects that drive greatest impact. Includes larger members of the agency teams
- This has now evolved to a stage where we are now looking at extending short term project based exposure through an exchange program for mid- level talent between Agency and PepsiCo

Over the years, this process has become a landmark event and a sort of a mini research industry meet. The senior leaders have applauded this process for the transparency and sharing of plans; and this has helped agencies make suitable resourcing plans and share calendars internally. It has helped drive retention of teams and helped instill pride in the account. This has also helped agency teams plan for new techniques and innovation as the teams have had visibility into future plans well in time. For the younger members of the team, this is a prestigious event and the awards ceremony something they look up to and work towards.

The Annual Agency Reviews and Meet at PepsiCo have become a landmark and something the agency teams wish other clients would emulate. Feedback from Key Account Managers of our leading agency partners:

“I can definitely say that 2016 was a different experience in many respects and it did take this process to a new level. Overall the 2016 agency review process was a great one and is bound to make a meaningful impact at both ends. Some of the things that really stood out for us were:

1. *Early sharing of feedback which gave us time to assimilate and prepare for a fruitful discussion on the review day*
2. *Action planning based on feedback was given more importance than back & forth on the feedback itself. It made the whole discussion forward looking and purposeful*

3. *Our teams were able to use the feedback for making concrete action plans for next year and think of ways in which we could work in a better manner and help PepsiCo business grow*

Annual Reviews have helped in a lot of ways especially on softer aspects. It helps focus on the areas that really matter and can make an impact. Gives you the context of what is important to business and therefore align the delivery to expectations. I think this streamlining of what we do has been most useful.

The recognition of best in class work has been very motivating as this is at a project level which specific members of team can identify with. It also helps in understanding priorities and therefore identifying areas where we can partner more / better."

Recognizing the power of the network, and the camaraderie that events like this bring along, we have moved onto to *Joint briefing and sharing sessions* for many of our critical Insights sessions. Quarterly reviews were the first to move to this, with joint briefing sessions and present back sessions with the wider audience encompassing the Retail tracking team, Brand health measurement, Innovation testing, Media and Planning partners. The power of the network to provide context, build on each other's storylines and jointly identify opportunities and threats has really helped the business benefit from the collective thinking of all key agency partners.

Seeing the value in this, the WPP plc (British multi-national advertising and public relations organization) lead teams have instituted this as a quarterly process wherein all the research, creative and media agencies that work with PepsiCo come together, chaired by a very senior leader at WPP, to jointly diagnose and brainstorm on ways to impact PepsiCo businesses.

The collective power is definitely superior to individual partners working in their silos, and once there is comfort in establishing the protocols and necessary changes, this clearly benefits all parties - client and agency teams alike.

One such change born from the agencies coming together is the agency teams proposing a *Key Account Manager*, who is a mid to senior level resource that spans visibility across a set of partner agencies, to best derive and provide solutions to clients. The KAM spends a significant portion of a workweek onsite engaging with multiple stakeholders. These engagements help alleviate the role of the agency team in everyday life of the stakeholders, provide first hand visibility into the concerns and issues facing business managers and enable wiring across the business.

In this way, collaboration has helped agencies alleviate their stance and role in the life of the client teams; while helping PepsiCo derive high quality engagement and thought leadership.

Taking this forward, PepsiCo Insights is now actively *engaging with Product, Insights and Marketing teams of other partners* **-** which Insights does not classically interact with - namely R&D partners like Flavour houses and Packaging teams; digital partners like Social Media Listening platforms and PR agencies; Logistics and Quality control partners, etc. In a traditional CPG format, Consumer Insights has limited interaction with these partners, and rarely proactively engages with them.

Recognizing that we would have to leverage existing relationships for newer synergies; the team has actively focused on inculcating these relationships now from a knowledge sharing basis, which has helped drive a number of Learning Sessions and helped feed into foundational insights.

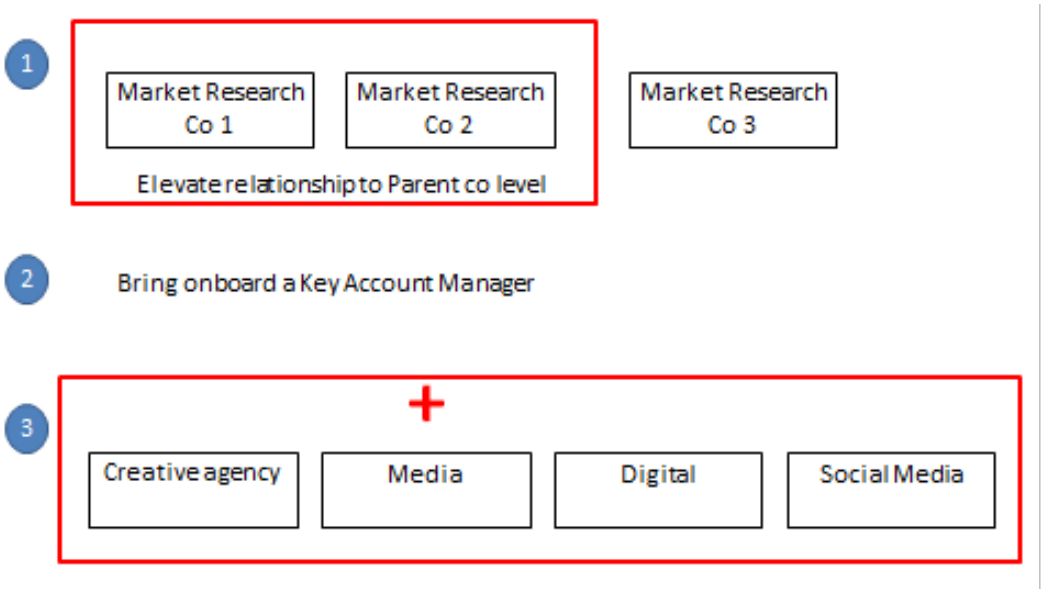
There is also a drive for greater collaboration between clients that PepsiCo has been championing through *Client roundtable session***s**, and a call to Agency teams to work towards Syndicated solutions for broad issues facing clients.

In its truest sense, investing in and leveraging the network has made it nearly a self- sufficient process, wherein in close partners are engaging with each other as much as they engage with PepsiCo towards impacting PepsiCo businesses. This is truly an enviable position to be in for any client, where partner agencies are leveraging and supporting ach other towards common goals for the client, vs. working in silos, or worse still competing with each other for client attention.

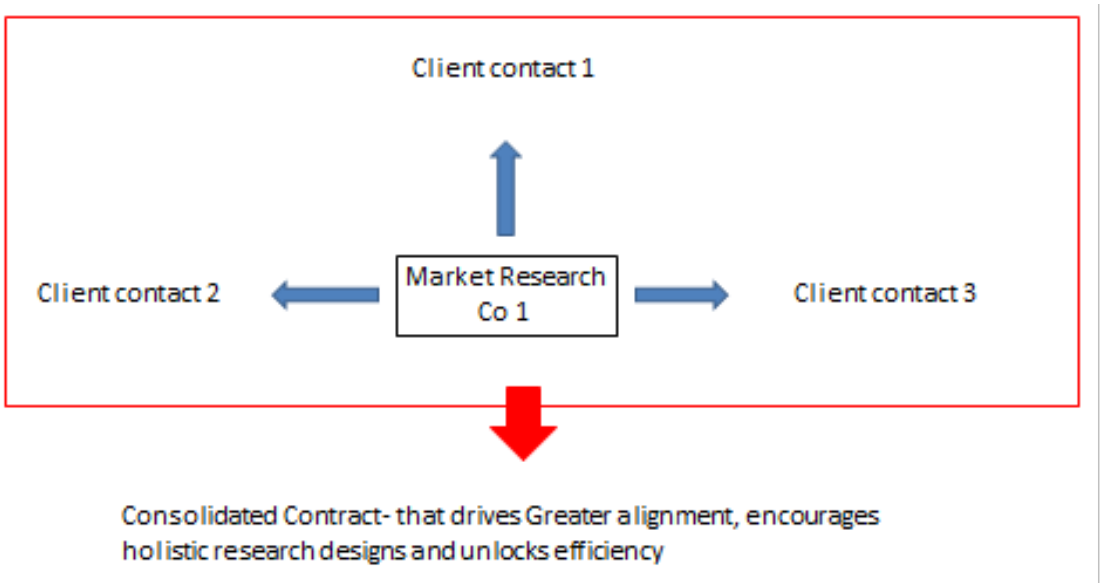
If you want to go fast, go alone. If you want to go far, go with others (African proverb)

There were four fundamental routes PepsiCo Insights took to affect the build out and use of the network.

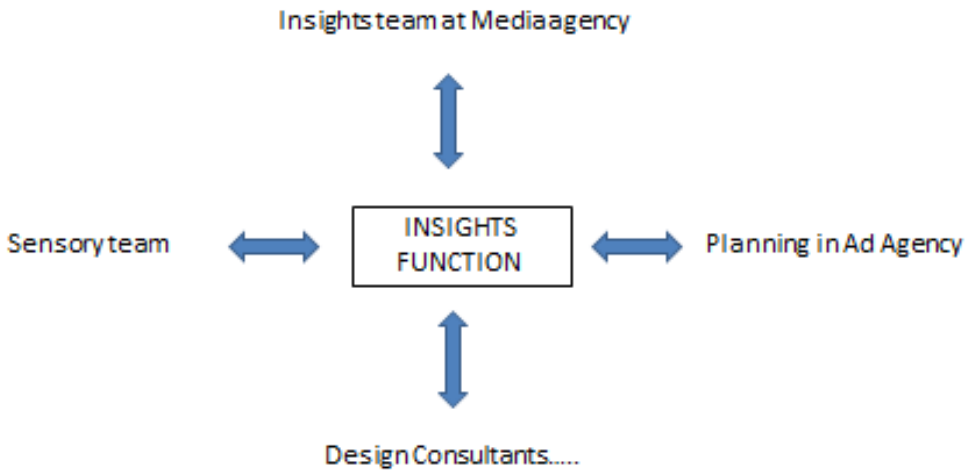
1. Building a holistic partner network across agencies



2. Consolidating multiple contracts across functions

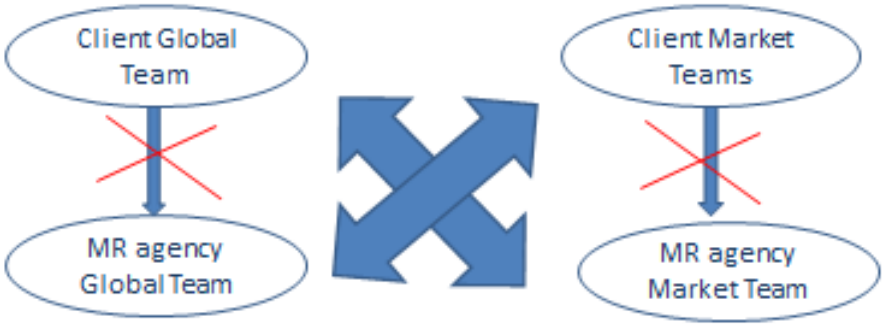


3. Building an "insights" network across systems



Bringing together Insights teams from across "Non –Related" organizations

4. Converging global and market wiring



From market wise interactions to a larger cross wiring between Global and Local

If you want one year of prosperity, grow grain. If you want 10 years of prosperity, grow trees. If you want 100 years of prosperity, grow people. (Chinese proverb)

By leveraging networks, the PepsiCo Insights team has been able to *integrate into the wider business*, now impacting Commercial, R&D, Quality and Human Resources over a period of merely two years. Therefore the most prominent impact that this drive has been able to generate is to widen the sphere of impact from Marketing alone to a much more holistic level. This was the desired outcome to start with, but to achieve this outcome with only marginal addition to resources, in such a short period of time, and without any disruption to existing relationships or workstream is commendable.

Having top-down and bottom-up connections have enabled *improvements in marketing execution*. Insights now leads the agenda for monthly functional meetings, has a dedicated 1-to-1 session with the CEO every month, partners HR on organizational and customer health initiatives and is represented at all key commercial meetings. This has helped ensure that both at strategic and tactical level there is insights representation in most forums.

One lens to view this impact is through the *shifts in responsibilities* bestowed to the Insights function which has elevated the role and impact of this team:

- *Reporting to action planning*:
  - From reporting market performance to Senior leadership to Setting targets for in market execution targets for Sales and Finance teams
  - From reporting share of shelf and in store execution to enabling teams to design in store experience
  - From reporting campaign performance to co-writing the creative brief

- *Project management to knowledge management*
  - A positive fallout of the insights team widening its stakeholder base is that Insights is now viewed as the one point contact for knowledge and information residing across the business – especially as a conduit between Marketing, R&D, Sales and Finance.
  - For example, the Insights team is now the custodian of reports like Market Visit and Consumer Visits (collected by Sales vertical), Pricing and Competitive activity charts (collected by Finance vertical).
  - Not only does this make the insight recommendation more holistic, it also makes Insights is the central nervous system of the organization

- *Brand building inputs to real time response management*
  - Integration of the Insights and Social Media support teams has meant we now have the capability to test real time, and more importantly respond real time to trending conversations - based on data backed knowledge of profile and motivators of the conversations on social media.
  - In one instance, we found negative commentary trolling on our flagship brand pertaining in response to a pricing change. Through the use of analytics we were able to profile the underlying causes of concern behind pricing- which had less to do with actual change in price, and more to do with shortage of supply of a different variant, which had made people resort to substitutes. We were able to zoom in on the micro markets and correct for this both through communication and in market activity; thereby averting it from snowballing into a larger issue.

- *Productivity*
  - We found productivity of nearly USD one million on a Quality measurement project. The Insights team in partnership with Operations, Quality Measurement and Sales helped liaise and align methodologies between two data platforms - the Sales team's Data Execution Governance system and the Quality team's Quality Check System.

For the eco- system, the turnaround was the most stark. To start with research agencies had some apprehension that this was largely a productivity drive, and the challenge of working and sharing information outside the agency was viewed as a big problem. But as the days progressed, partners realized that the benefits far outstrip the challenges. Some of the clear *advantages voiced by the eco-system* are as follows:

1. Greater transparency and visibility into client plans
2. Therefore an ability to plan internal resourcing, innovation and calendar
3. Wiring across levels and teams; including access to Global colleagues
4. More empowered Insights function that means greater exposure for research partners
5. Ability to provide holistic solutions and stories by partnering with non- compete agencies with different skillsets
6. Greater efficiency and productivity as a result of greater visibility
7. Overall improvement in business value
8. Thriving and engaged teams, lower attrition levels
9. Lead account for proactively experimenting and trying new tools and processes
10. Access to learning and shared resources

What is also important to note is that this is a evolutionary process and therefore the more you participate in it, the openness to failure is higher. Along the way, some processes and tools will have to change, and change management is never easy. But by building out an eco system, the power of the collective supersedes these minor hiccups.

To really ensure that this is a win-win for both parties, and to bring this journey to life, there are five critical pillars

- Alignment of goals – in the absence of which this will not be a symbiotic relationship
- Setting and charting out ways of working
- Balancing the short term priorities with the long term imperatives
- Senior management commitment to the new structure and processes
- Leveraging technology to automate processes and upskilling teams on digital skills to be able to work with larger location agnostic teams in real time

This is a journey and the road to this evolves every day. The path becomes much clearer because we chose to leverage our eco system along the way. In sum, I would like to share one of my favorite reads:

"What makes networking work is that it sets up win-win situations in which all parties involved get to take something home. Networking is a sharing process. Until you understand that, you won't have much of a network." (Earl G. Graves, Sr.)

The Author

Madhumita Chakraborty is Associate Director- Consumer Insights, PepsiCo, India.

Streamlining the Innovation Process to Maximize Efficiency

Cross-stakeholder participation and involvement from the get-go!

Allen Cai • Ching Ling Ng

Introduction

Few of us would argue that innovation is a buzzword. It has and continues to be used in industries from education to healthcare to fashion to consumer goods, all around the globe, and can mean anything from a new idea to a new design, a change in business model to a change in thinking. Likewise, Innovation has been floating around in the research industry for some time.

The type of innovation we want to shed light on in this paper is about implanting and encouraging the creative spirit of innovation to propel client teams towards the actual innovation of the product and activation by themselves.

What is innovation?

Innovation is about developing the products and services that the market needs tomorrow, and is driven by the need for convenience, not by technology. Finding a 'cool new idea" while it may be interesting and exciting, has no value unless the idea solves a real problem. Once the consumer opportunity has been found, the problem solving is where most people recognize the innovation as taking place. Breakthrough innovation comes from finding radical solutions. This is where connecting with a product or process from a totally different environment often leads to that 'aha' moment.

Contrary to popular belief, new ideas emerge most frequently as the result of collective knowledge, and typically do not come from a lone 'genius'. Consequently, for successful innovation to occur, the internal and external networks of an organization have to be well developed (Marill, 2015). Innovation is therefore a team effort whereby a well-balanced, participative group of people produce a plan to sell the creative concept and follow it through to implementation (von Stamm, 2002).

Looking at the retail racks in China, one will be surprised at the number of variations of Lays, Oreos, and new brands appearing on the shelves faster than you can name all the countries of the world. However, not all brands and products succeed. The vast number of product innovations in the China market boils down to some key factors.

- *Agility of local players in introducing product innovation to the market at a very fast speed*
  Faced with an increasingly complex shopper world where e-commerce plays a big role in China, and the role of traditional shopping channels are evolving, local players are quick in adjusting their portfolio and introducing new 'innovations' from packaging to flavor, to totally new products. Local players are quick in sourcing, and bringing their innovation to market with a much quicker speed. Having worked with some local clients, we also learnt that the moment they feel that it is a good idea, they will put the investment behind it to bring it to market at the shortest time possible.

  For a global company such as Unilever, there are many processes involved in getting an innovation to market. Therefore, one of the objectives of this streamlined process is about coming up with as many ideas as possible for short-term quick turnaround to long-term investments that require various stakeholder buy-ins. This was to ensure that we will have foresight on consumer needs so that we are creating future-proofed innovation that would be able to withstand the demands of the various rounds of concept testing.

- *High investment into big activation campaigns that achieve higher resonance with the domestic audience*
  With a totally different world of activation compared to global markets, Chinese brands also seemed to do much better in terms of activation campaigns. Take “Student Xiaoming”, a beverage drink that was introduced in 2015 by Tung-Yi, for instance. From the get-go, it had a targeted positioning for the Chinese millennials who yearned after nostalgia. Having to push its way in a long-established category of ready to drink teas in China, they went after a ‘niche’ positioning, supported by a highly differentiating campaign that effectively cut through the clutter of the advertising. It also tapped into social media to activate and engage its target audience with sharp messaging that talked the talk of their target audience, thereby achieving over RMB 500 million within half a year from launch.

Figure 1. “Student Xiaoming” beverage campaign activation in China

Typical stages of a process of innovation via http://www.autenburgfin.de/fileadmin/_processed_/csm_marketing-phasenverlauf-innovation_e_0bbe8c32ef.gif

While this is just a case study of one brand, wherever you go in China, especially Tier 1 and 1.5 cities, it is no longer surprising to find yourself bombarded with advertising messages from the moment you leave your house.

Created for China, in China, by China

The mission we were handed was to supercharge China-specific innovations that were created for the market, by stakeholders in the market, with regional and global teams being a part of the process to truly bridge the cultural gap in the global innovation process in order to create innovations that are truly meaningful and impactful for the local market.

Redding and Drew (2016) also point towards the theory that while innovation could be culture agnostic, cultural sensitivity is also needed in innovation. Hence while most of the participants and the impact of this innovation process would likely be produced by the China team, regional and global team members were also included to ensure that learnings can be cross-shared and the collaboration process would empower the China market to take action locally in order to succeed in the market, rather than waiting for the different layers of alignment across time zones and culture.

The role of market research in innovation

Traditionally, the role of market research has been more about exploring, understanding, tracking, and of course, reporting. However, with digital innovation in research, consumers have evolved from subjects being closely watched under a microscope of qualitative, quantitative, neuroscience and other emerging tools of market research, to partners, collaborators and co-creators to help brands succeed and grow in today's marketplace.

When you ask any client with a sizeable research budget how many reports they receive a year, chances are it's going to be anywhere from 20 to 200, ranging from U&A study, generational study, countless concept tests, shopper studies, and the list just goes on. It's hard to stand out among the many reports that are received from the big agency players to the smaller boutiques that offer a good run for the money with their innovative, non-traditional approaches, and promised sharper insights.

It is not just about the amount of investment and data that clients gathered, it's about what is being done with the information to make it useful for the purpose of keeping the client business relevant, and future proofing the client business.

Having the market research function be the lead-voice of the innovation process and being the facilitator to streamline the process was a critical factor for success. The market researchers are knowledgeable about trends, consumers and their needs, and moreover are able to distill the insights into easily digestible information pieces. This brings all stakeholders to a common starting point and facilitate cross business unit communication and collaboration in order to make the outcome actionable and impactful for the business.

In traditional processes, market research would have played the role of identifying unmet needs through qualitative and quantitative research, perhaps even including Semiotics to identify potential opportunity areas. Once concepts were developed, there would be various research that incorporated concept optimization, testing, prototype placement before any product would even be launched.

This was a process that typically takes three to five years to bring any innovation to market. However with the advent of digitalization, the need for agility – especially for a well-established company which had a stringent innovation process and check steps – was important to keep up with the fast changing consumer landscape and the quick agility and adaptation of smaller brands and corporations.

What goes into a typical innovation process

There are processes that are in place especially at big companies to bring innovation to market. It takes rigorous planning, going through the upper ladders of approval, opportunity sizing, market testing before any ideas would ever see the light of day.

- Market assessment
- Internal assessment
- Idea generation
- Internal alignment
- Development
- Fine-tuning
- Evaluation
- Positioning

While these don't look like a really long process, it takes years for products to go through all the stages in order to become reality.

What we tried to achieve in this new innovation process is to cut down the time of innovation with leaner processes in order to get ideas to market in a shorter turnaround.

Garnering senior support and participation from planning

Empowered by senior support and participation from local, regional and global leads, the team was able to set a fixed date for the innovation session. With senior leadership involved, stakeholders were brought on board early on in the process in order to ensure the active participation of the internal client participants across seniority, across function which served as an enthusiastic enabler for the possibilities of the workshop.

Rather than bringing in consumers to co-create, the innovation process was streamlined to just bringing in stakeholders that know about the business issue, processes, pipeline and feasibility to make the output of this process easily activated.

Choosing the right facilitator, one that can think on their feet

With a diligent system in place for qualitative moderation assessments, Unilever typically only works with accredited moderators. Having worked with many moderators across agencies, the insight team was quick to notice that it required more than just qualitative moderation skills to facilitate a workshop of this scale that needed to both bridge cultural and language barriers.

It was essential to select a facilitator that could think on the spot, be credible and genuinely interested in the success of the client's business in order to run a workshop that involved junior to senior participants in the same room. The agency needed to be vested in the success.

Distilling the key insights from the knowledge bank to give all stakeholders a good starting point

To bridge the cultural and language barriers, and furthermore give true insights to the future generation of consumers, the Unilever insights team distilled learnings from generational studies, category usage and attitude studies, digital behaviors and trends. All of these were presented as bite-sized and easily digestible information in order to facilitate easy absorption by a room of 40 client participants.

Inspiring stakeholders with education, trends and interest

Beyond just hearing from the internal insights team on trends, consumers, usage and attitudes, and things that might be put in the bucket of 'I already know that', it was also essential to inspire stakeholders with experts from adjacent categories with new knowledge that brings a new light to the way they view their category.

Letting stakeholders immerse on their own with interesting tasks to fulfill

Rather than traditional approaches of market visits, and consumer ethnography conducted by a qualitative moderator, we inspired the stakeholders to be involved and responsible for their own learnings. We took away the safety net with blessings from the client, and let them go out on missions that were both educational as well as exciting. We even let them roam into homes of consumers to ask the questions that they really wanted to know, to experience the usage journey of real-life consumers and not just read it off a beautifully designed PowerPoint slide with a picture, a process, and some verbatims. They had to be immersed by themselves without the safety net of an agency.

Figure 2. The immersion kit was designed to be a mix of challenging tasks and tips on how to engage with consumers

The immersion kit developed for Unilever

Bringing it all together to kick-start the creation session

Equipped with knowledge from insights, experts and their own immersions, we started the creation session by giving a quick recap of what they have been through, and what they have learnt with a gallery-style walk through of learnings, coupled with a quick video recap from their immersions where they could interact by sending comments which were real-time showed on the screen. This helped to capture their attention and fuel it with excitement for the start of the day.

Rounds of creation outside their everyday work before bringing it back

With some stakeholders, having worked in the business for a long period of time, and schooled in the ways of working and boundaries of the brand, we sought ways to get them to, although it is a cliché, 'break out of the box'. We designed ideation and creation rounds that brought them into parallel universe, in order to develop completely new ideas that were not existing innovation ideas already in the pipeline. This inspired them to think and walk in the consumer shoes, and other roles which in each step challenged their thinking. With these out-of-the-box ideas, the success factor was still to be relevant to the client's brand, therefore a final round of refinement and presentation was made to bring it back to the World of Unilever.

In order to facilitate the rounds of testing, visualizers were also on hand to create full concepts including a visualization of it to get it to round quickly through the established idea screen.

Tapping onto the post-workshop excitement to further engage stakeholders and keeping them in the loop

While the creation day ended with an appreciation speech by a senior stakeholder, it was still critical to "strike while the iron is hot". Hence we tapped onto the post-workshop excitement with a visual recap to thank and appreciate all participants for their involvement and keep them excited and engaged for when the real work of innovation starts.

This was achieved through a three-day process. It required active involvement from stakeholders which was made possible by the senior endorsement from the get-go.

Some of the ideas have already been brought to market, while the bigger product innovation ideas have gone into testing and development. Instead of having a long drawn out innovation process from understanding consumer needs all over again to the various stages of innovation, this process has streamlined the innovation process to make it more tangible and activation-oriented. While we are not saying that all companies should bypass the traditional model, this has helped us achieve our goal of streamlining the innovation process for a client that is already very knowledgeable and respected in the world of insights.

A streamlined process such as this one brings a more important role to the client side and unconventional challenges for a traditional research agency. However, it also brings a heightened sense of involvement with consumer insights as the core of innovation processes to bring consumer-centric innovations to market.

The more the merrier, yet only the best will come to pass

One of the briefs of this workshop is not only producing good quality ideas that could get bogged down by participant's passions and wishes, but rather, to provide a good quantity of ideas.

From the drawing board, we set out to achieve quantity of ideas. With over three rounds of ideation, we had over 100 ideas that came from the internal team participating in the workshop. Some of these ideas have been in the heads of some of the team members for a long time, yet they had little places to voice these ideas, for others, they didn't have the time and resource to put their ideas to paper. This process offered the avenue for all of the team to have a voice, participate and input towards the future innovation of the brand.

We've learnt from past innovation projects that without a sizeable amount of ideas worked out in concept formats, it is hard to brief agencies, and it is harder to pass through the rigorous rounds of consumer testing in order for the innovation ideas to be able to be developed and go to market.

Conclusion: We impact lives

In past rounds of innovation concept testing, we typically would only see three or four ideas getting passed into the next rounds. However, with this supercharged innovation approach, we got 10 ideas into the next stages of development. As this is a recent project, and it is still not yet introduced to the market, sit tight, and wait to experience the ideas that came out of this workshop.

Too often innovation programmes can be hampered in two main ways. Either the group, or even just a few powerful voices within it, adopt a 'yes but' attitude, stifling progress by identifying potential minor flaws or drawbacks which they put forward as a reason for dropping various ideas. Often this attitude stems from a lack of confidence, and a risk-averse attitude so encouraging stakeholders early on to take leaps of faith can pay dividends here. If the 'yes but' attitude prevails, you can go from having 30 viable ideas to just a couple; and these couple of ideas will usually be amongst the 'safest' and least inspiring.

The other way, at the other end of the spectrum, that the process fails to yield results is when the group can't let go. Someone's proposed an idea; they believe in it; and will defend it to the end. They begin to see any criticism of their idea as a criticism of their own creativity and commercial awareness. This also stems from the fact that innovation, as noted, has been a buzz word that everyone talks about; so everyone knows is important; but few know how to do it well. As a result these innovation sessions can be seen as a 'test' of how up to date and able stakeholders are, hence the potential tendency to never let their idea die. That's why the critical importance of setting the scene and putting people at ease early on cannot be emphasized enough. It is even worth taking attendees through the above, and even citing the Google mantra that failure is good for success.

What this means to researchers is to keep up with the dynamic changes in the market, and stay abreast of developments in order to enhance:

- Agility in bringing together and distilling useful insights for the purpose of the session
- Collaborative spirit and approach to bringing internal stakeholders together to facilitate meaningful conversations
- Engagement and open channels to participate and raise questions
- Empowering internal clients to walk in the shoes of consumers and recall their identity as a consumer
- Inspire consumer-centricity in innovation and ideation
- Rapid switch on idea screening
- Driving force in bringing innovation forward

Bibliography

Marill, P. (2015). Innovation Never Stops: Innovation Generation – The Culture, Process, and Strategy. ASQ Quality Press.

von Stamm, B. (2002). The Innovation Wave: Meeting the Corporate Challenge. John Wiley & Sons.

Redding, G and Drew, A. (2016). Seven Attributes of the Most Innovative Cultures, Journal of Interdisciplinary Economics , University of Newcastle Business School.

The Authors

Allen Cai is Senior CMI Manager Country & Customer Development China, Unilever, China.

Chingling Ng is Research Director, Labbrand, China.

PART 2
Client-led Discoveries in Consumer Insights

Sonification

New way of engaging with consumers in scientific product communication

Noriko Nakano • Eri Yamashita • Sono Hasuka • Jung Yeon Nam • Mariko Nomura • Damien Velleman

Introduction

Product knowledge and the prior product use experience are two of the key factors influencing consumers when they are deciding to purchase a product. According to Court et al (2009), consumers have four primary phases of decision making for purchasing: initial consideration, active evaluation, or the process of researching potential purchases; closure, when consumers buy brands; and post-purchase, when consumers experience them. In the initial consideration phase, communication about the product use experience is critical for consumers to move on to the next process to make the purchasing decisions.

Business challenge – How to communicate sensorial benefits of cosmetics triggering consumer interest?

Communicating the prior product use experience is a key challenge related to cosmetic products, and sensorial benefits are some of the key factors for purchasing decisions. Sensorial benefits of cosmetics are mainly communicated via visual images in advertisements. Song (2010) indicates that of the five human senses, visual information is the primary influence on how we analyze things around us.

However, visual senses alone have a limited scope in communicating product benefits to consumers. In fact, we try to use our multiple senses to better understand the products that we encounter in our daily life. Song (2010) argues that when we touch something that we don't see, we try to determine what we are touching with the sound, smell, and taste of the object.

Sounds and music for enhancing the consumer experience through the sense of sound

Sound (and touch, as well) is a human sense that is associated with physiological and psychological stimuli. The study by Brown et al (2011) shows that blindfolded participants are able to use a visual-to-auditory sensory substitution device in naturalistic use of experiments to locate and identify objects. Sonification is one of the emerging approaches which could exploit the sense of sound in human beings.

Sonification model

What is Sonification?

Sonification is a communication approach that uses audio waves (e.g., sounds and/or instrumental music) to transfer information. According to Kramer et al (1999), Sonification is defined as "a transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation". Sonification has also been successfully used in data analysis and exploration tasks, proving fruitful where visual techniques have not, such as sensory substitution for visually impaired users (Kramer et al 2010).

Possibilities of Sonification for communicating sensorial benefits

Sonification could be regarded as useful in media communication for products which require a relatively long period of time for consumers to feel and sense the product experience. One example is alcohol-related risks caused by high alcohol consumption. According to Walus et al (2016), their study shows that music and Sonification could help to convey alcohol-related risks to young people.

However, Sonification is still a new arena for communicating sensorial benefits to consumers. Currently, the main promotion media for communicating brands and products are visual messages such as images or words. This is the reason why we aimed at exploring Sonification as a way of translating through sounds and music the intimate feeling(s) of consumers' experience using a cosmetic product.

Sonification prototypes for hair treatments

We explored the Sonification approach with hair conditioner in order to relate the perceived feelings (by touch) brought by hair conditioners to a sound or music generated by a frictional instrument that transforms physical signals into audible waves.

The Sonification prototypes were developed by taking two steps: i) assess the ability (sensitivity) of currently available instruments to record the frictional state of differently damaged hair swatches and ii) develop an algorithm that transforms the recorded frictional forces into sound which can be perceived by humans (i.e. between 20 Hz to 20 KHz) (M. Nomura et al 2016). The algorithm was developed by transforming hair friction data into friction signals and music. (See figures 1 – 3).

Figure 1. Process of raw data conversion into sound

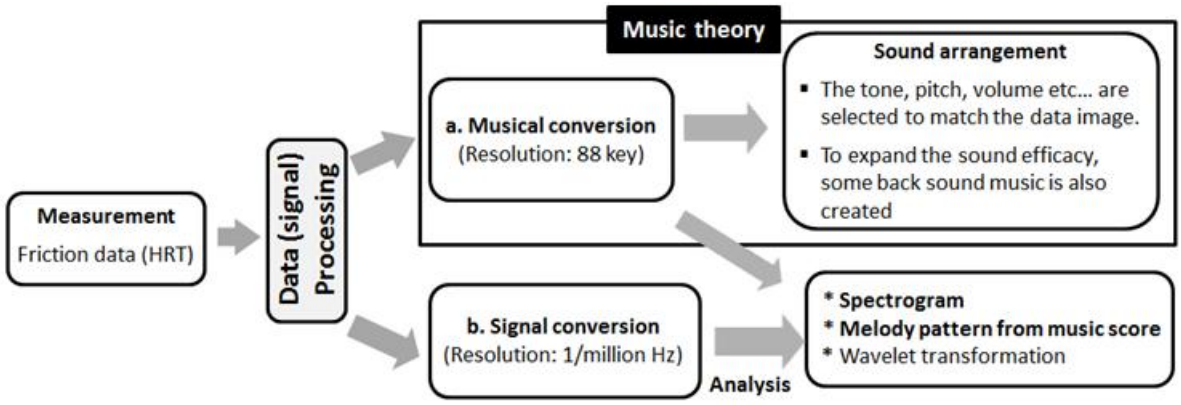
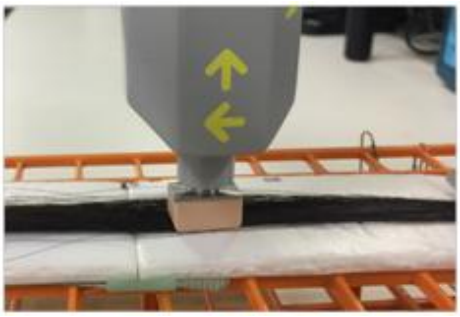
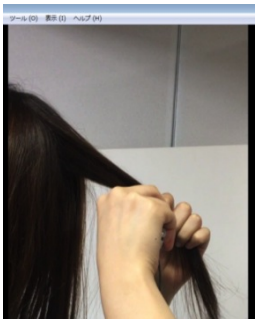


Figure 2. Frictional force evaluation (Ft/Fn) and Photos (Left and Right): The TL 701 Handy Rub Tester device for the friction measurement (left) and its on-head measurement of hair friction (right)



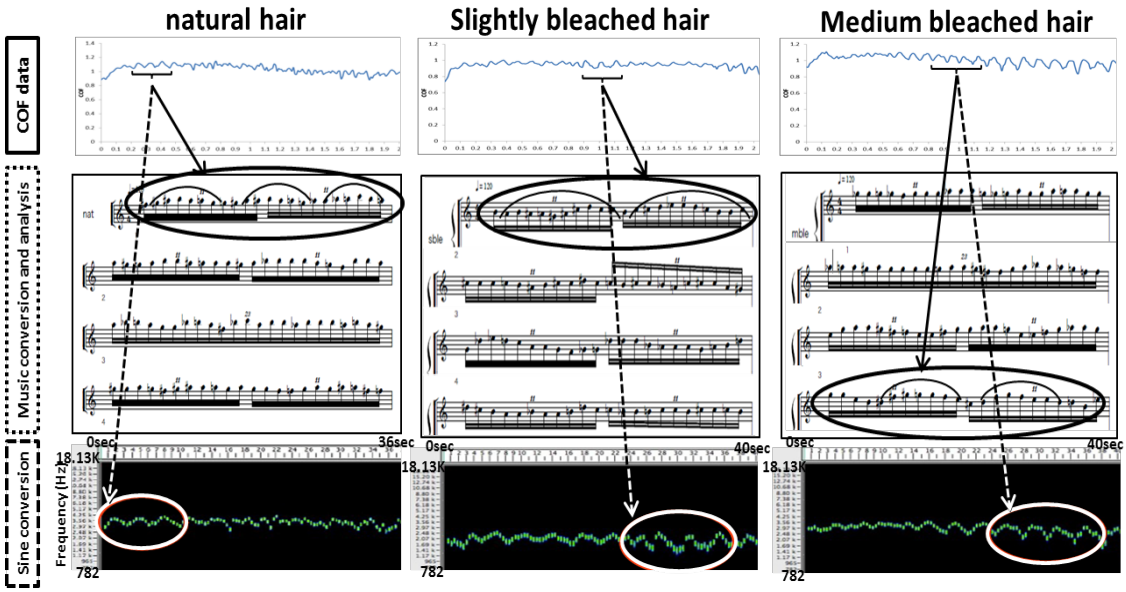


Can Sonification emulate the sensorial experience with cosmetics for consumers?

This is the most critical question for Sonification to prove its effectiveness for communicating the sensorial benefits of cosmetics. The following questions should be answered to confirm the effectiveness of Sonication:

- Can consumers perceive the sensorial characteristics of cosmetics in line with the technical product performance?
- What added value(s) could Sonification bring to product communication compared to other approaches, such as those with only visuals?
- Can Sonification work for everyone regardless of consumer nationality?

Figure 3. Musical analysis on the music score and comparison to the spectogram



Case Studies: Consumer perceptions on hair treatments in Japan and Korea

In order to answer the above questions, we conducted a series of preliminary consumer perception studies on Sonification with hair treatments in Japan and Korea where the hair treatment markets are well developed.

Tested products: Haircare treatments

In this study, we used haircare treatments for the following reasons:

1. Haircare treatment was the most suitable product category to start with the Sonification prototype productions by collecting various technical data using various hair swatches.

2. Experiments on Sonification using hair treatments allowed us to explore various application opportunities including Sonification as part of media communication and Sonification at hair salons.

Laboratory assessment case: Initial consumer perception study in Japan

As the initial consumer study, we conducted one-on-one interviews with Japanese consumers to: i) assess if the consumer is able to understand different hair conditions via the sound differences, ii) examine consumer interest in the Sonification approach, and iii) identify any areas for improvements.

Method

In this study, we recruited 11 Japanese females aged 20-49 years old with medium to very long hair to complete the following steps and have a one-on-one interview at our consumer interview facility in Tokyo.

Assessment 1: Can consumers perceive hair condition by just listening and watching Sonification without touching hair?

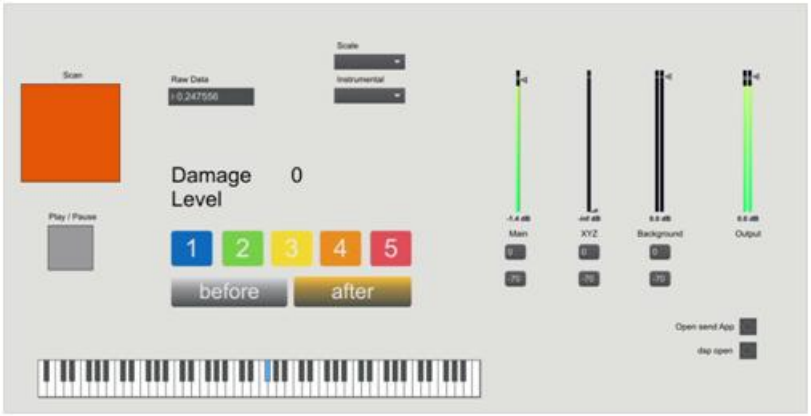
Each respondent listened to the following four types of sounds and simultaneously watched corresponding videos showing hair friction measurement. Of note, each sound sample was presented in a randomized order without giving information about the hair condition.

Figure 4. Types of sounds presented to consumers during the study

Type	Description of Sonification Prototype		
	Hair Condition	Hair Friction Measure*	Music Instrument
1	Before Hair Treatment	0.494	Harp
2	After Hair Treatment	0.303	Harp
3	Bleached Hair (damaged hair) with Harp	0.543	Harp
4	Placevo (Music only)	-	-

Hair Friction Measure: Hair co-efficient friction) value. Higher value means higher hair friction*

Figure 5. Visual image of hair friction corresponding to each hair condition, shown simultaneously with sounds and music



Remark: Type 4 (placevo) did not have the visual image, but music only.

Assessment 2: Can Sonification help consumers perceive hair conditions by touching, in comparison with touching hair without listening and watching Sonification?
In this assessment each respondent touched a hair swatch in the following conditions:

Figure 6. Types of sounds presented to consumers during the study

Type		Hair Condition	Hair Friction Measure*	Music Instrument
a.	With Sonificaiton (Harp)	After Hair Treatment	0.303	Harp
a'.	With Sonificaiton (Marimba)	After Hair Treatment	0.303	Marimba
b.	Without Sonification	After Hair Treatment	0.303	-

Remark: Each sound sample of a and a' was presented in randomized order.

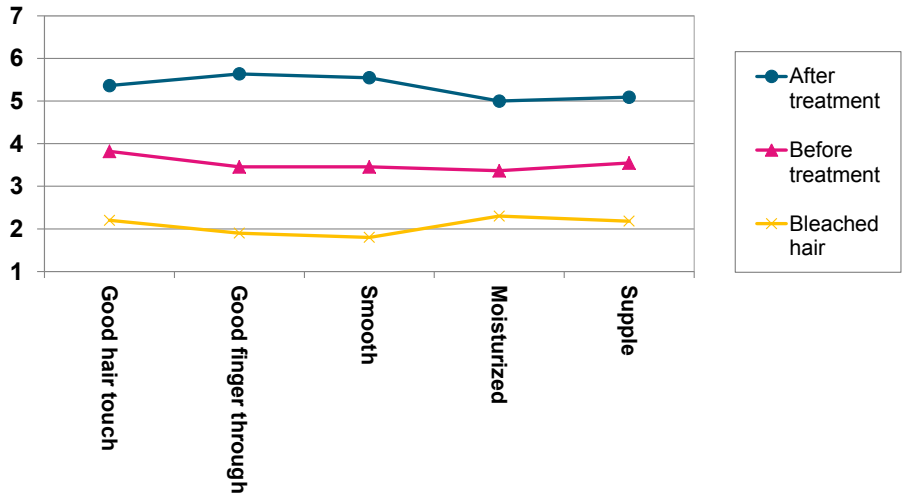
In each of the two above assessments, each respondent evaluated her hair conditions using the following attributes on a 7-point scale in addition to open-ended comments.

- Hair conditions – "good hair touch", "good finger through", "hair smoothness", "hair moisturization", and "hair suppleness"

Results

Consumers associated Sonification with perceived hair - Overall, consumers were able to associate the hair quality with Sonification prototypes by listening and watching the Sonification. Perceived hair quality was in-line with hair friction measurement and its corresponding sound/music.

Figure 7. Attribute rating of hair conditions assessed by respondents (without touching hair)



Consumer reactions to sonification prototypes:

- *Importance of visual data presentation together with sound:* While some respondents found benefits in the sound itself, the majority of the respondents preferred to look at the numerical data on the demo screen together with the sound. They thought that numerical data was convincing them to understand their hair condition.
- *Preference of musical instrument – harp vs. marimba*: Of the two musical instruments, the harp was received as a more pleasant sound than the marimba, mainly because of their familiarity with the harp sound in their daily life.
- *Comparison of with vs. without Sonification*: There was no difference in hair touch perception with and without listening to the sound: neither enhanced the perception of the hair touch. This was probably due to the in-vitro condition (e.g., touching hair swatches and not touching their own hair) which we need to explore more in real life conditions.

Improvements to be made for the next round of studies

The study found that the following improvements should be made as the next step:

- *Test method*: Conduct the Sonification in the real hair salon condition with specified target consumers because the hair quality perception was not enhanced by Sonification in this test, probably due to the in-vitro condition (e.g., touching hair swatches).
- *Sonification prototype*: Make the demo screen of the software more visually attractive to increase consumer interest.

Hair salon case in Japan: Consumer perception study in Japan

Having made the improvements identified after the initial study, we conducted the second study on Sonification with the following objectives: i) to explore potential benefits and the impact of the Sonification approach on hair diagnosis, particularly by comparing consumer reaction to hair diagnosis both with and without Sonification, ii) to obtain consumer expectations for hair diagnosis with the Sonification approach, and iii) to identify improvement opportunities in developing hair diagnosis with the Sonification approach.

Method

We conducted the following four focus group interviews among Japanese females who have hair damage concerns:

Figure 8. Focus group compositions

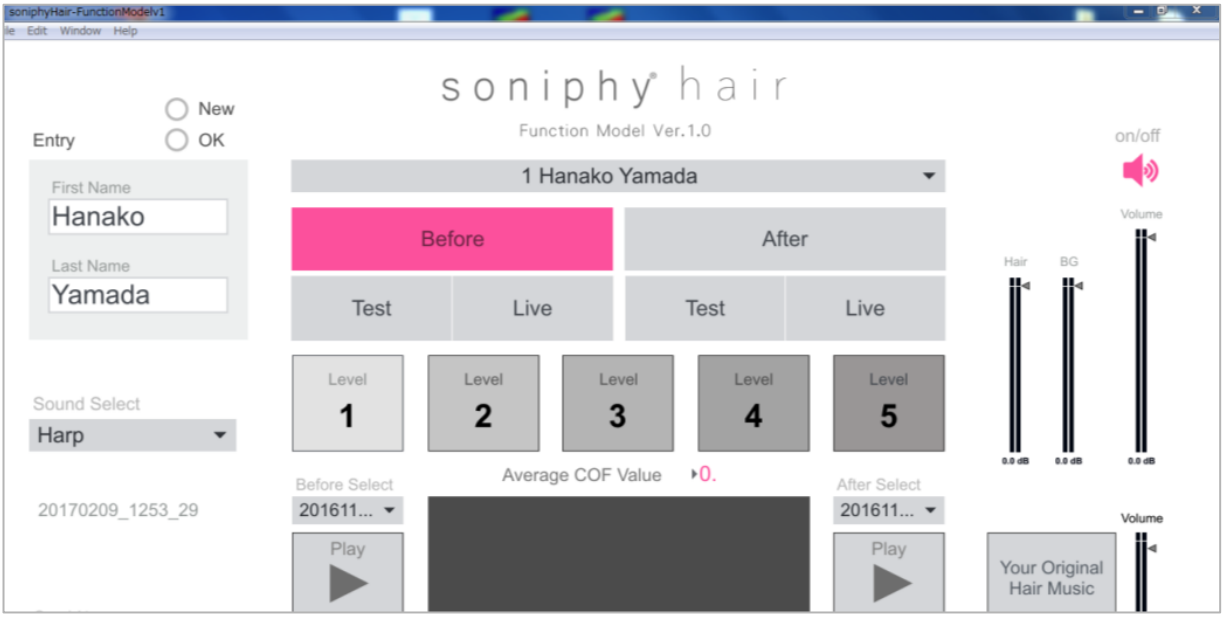
Group	Age Range	Tested Condition
A	25 - 34 years old	With Sonification
B		Without Sonification
C	40 - 49 years old	With Sonification
D		Without Sonification

The study was conducted in the flow illustrated in figure 9.

Figure 9. Test flow

Using the standardized gesture by stylist

Figure 10. Displayed visual on the screen of Sonification



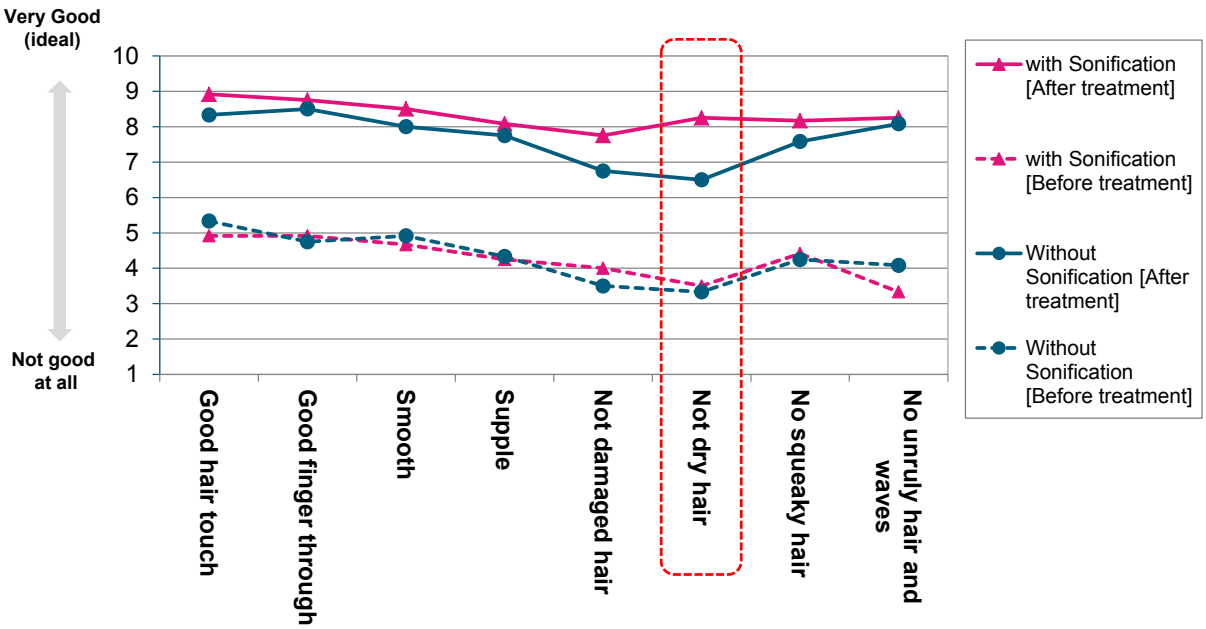
Results

Consumer reaction to Sonification at hair salon:

- Hair diagnosis with Sonification at a salon can be a special experience for consumers, as they were relaxed listening to the sounds and enjoyed experiencing hair conditions with sound. Such emotional benefits were the main benefits of Sonification, while numerical data tended to be referred to in order to better understand hair conditions.

- On the other hand, hair moisturization appears to be better perceived with visual hair diagnosis with Sonification versus with only visual diagnosis (without Sonification). There are two potential causes for the better evaluation: i)moisturization benefits were enhanced more by Sonification than by the technical facts, or ii)there is a gap between consumers' perceived moisturization benefits and the technical data.

Figure 11. Self-assessments of hair conditions Sonification



Note: Data show the mean score of each group

Further improvement opportunities on Sonification:

- While the melodic arrangement of Sonification sound tested in this study was pleasant to listen to, consumers expected to receive hair diagnosis with multiple measurements of hair conditions, such as moisture level, scalp conditions, and hair conditions, and they expect detailed consultation from a stylist in order to select a suitable salon item and to improve their hair conditions.

Hair salon study in Korea: Consumer perception study in Korea

In parallel to the consumer study in Japan, we conducted a consumer study on Sonification among Korean consumers to gauge Korean consumers' reactions toward this new approach and to identify improvement areas by understanding the overall performance of Sonification. More specifically, the study was aiming to: i) see Korean consumers' acceptance of Sonification, and ii) explore how Sonification could impact the hair treatment experience at the salon for consumers.

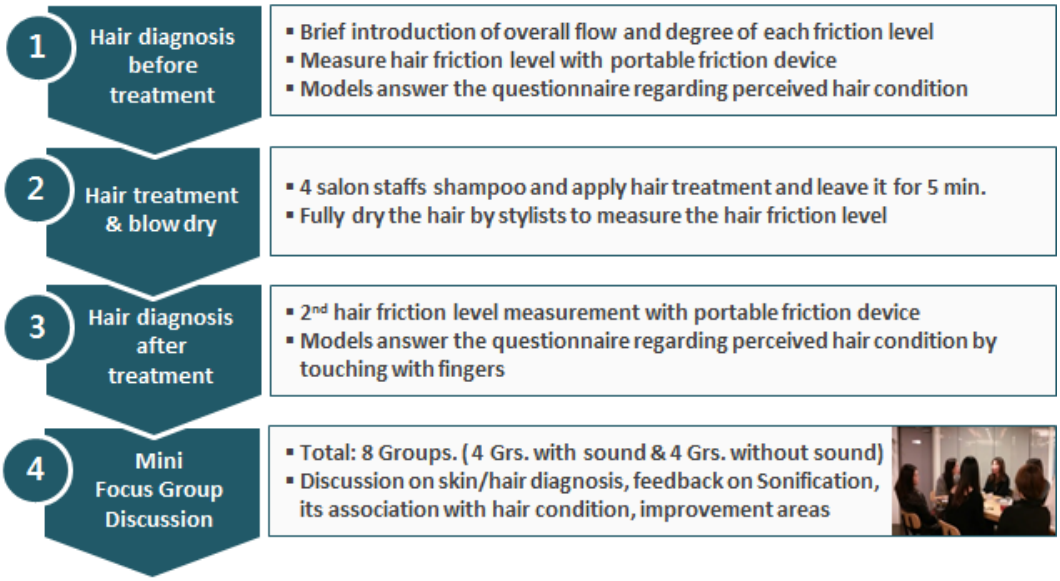
Method

We conducted the study with 32 consumers in the follow process, which was the same as that in Japan:

Figure 12. Focus group compositions

Group	Age Range	Tested Condition
A	25 - 34 years old	With Sonification
B		Without Sonification
C	40 - 49 years old	With Sonification
D		Without Sonification

Figure 13. Test flow



Results

Acceptance of Sonification among Korean consumers

- Overall, Korean consumers exhibited their interest in Sonification as a new device since Korean consumers do not have much experience with receiving hair condition diagnosis (other than the scalp). The specific perceived values are:

    - *Visual and Sonification added the technical data*: The respondents believed that Sonification has the added benefit of easily determining the current hair damage level in a short amount of time.
    - *Sound and music provide added aesthetic efficacy*: Consumers comprehend the hair condition diagnosis results shown via numerical scores and also appreciate having the results expressed in audio format.

Improvement opportunities for Sonification

- There were no clear differences in sound and music between before and after treatment. The difference in sound and music should be more distinct. Consumers need to get a full picture of hair conditions other than hair friction. Although the hair damage level is determined, models want to receive more information, such as data on hair thickness, after the diagnosis.

Overall key findings

Our findings confirm that sound and/or music, in addition to visual images, can be a key medium for communicating product benefits, which are by nature subtle and perceived during a variable period of time. Furthermore, sound and music allow consumers to closely synchronize their physical perception with their own emotions (positive or negative moods).

Perceived benefits of Sonification

- *Sonification is received positively among Japanese and Korean consumers in the context of hair conditioners*: Overall, consumers were able to associate the hair condition with Sonification (amplified sounds and music). Perceived hair quality was in line with the technical measurement of hair condition. The studies in Japan and Korea provide the possibilities for expansion of Sonification approaches in multiple countries, at least Japan and Korea.

- *Sonification enhances the emotional experience with products*. Listening to the sound and music in the salon gave consumers a pleasant feeling compared to listening to the sound and music in an interview room. The environment of listening to the Sonification is one of key factors for consumers to amplify their product experience.

Improvement opportunities

- *Multiple characterizations of hair conditions with Sonification*: In addition to hair friction, other characteristics such as hair softness and shine should be measured and communicated to a consumer to give the full range of diagnosis on her or his hair conditions.

- *Full engagement of Sonification with technical data*: In addition to sounds and music, associating images enhanced the perception of hair texture. While some respondents found the sound greatly convincing, the majority of the respondents preferred to look at the numerical data while listening to the sound. This will allow Sonification to avoid any exaggerations of perceived benefits through product communication.

- *Customize the Sonification approach by target*: The studies suggest that one Sonification approach does not satisfy all consumers. Therefore, it is important to customize the Sonification tool for the target in order to implement the approach in the real market place.

Implications to brands and products communications

Sonification includes sensorial benefits communication in advertisements, scientific communication to and education of consumers regarding the sensorial experience of product use, and the amplification of the product use experience with visuals and sounds.

Once Sonification is confirmed as the communication tool for sensorial experience with products, they can be applied to prior-use communication in advertisements and even sharing sensorial experience of users each other in the social network communities. To be more specific, sensorial experience of products such as hair smoothness before and after product use can be communicated with consumers through the digital media, which integrate Sonification and visual images so that the consumers more clearly understand the product use experience even before purchasing the product.

Furthermore, we also expect Sonification has the opportunity for Personalized Cosmetic Advisory Services. In addition to personal diagnosis of hair or skin conditions, we could add sound and music through Sonification in order for the customer to understand her or his hair or skin conditions not only logically, but also emotionally. We anticipate that Sonification could offer pleasant experiences to the customer, as well.

Future studies

As the next steps, we plan to further improve the Sonification approaches based on the learning from the consumer studies in Japan and Korea. More specifically, we will improve the Sonification tool to be able to characterize hair conditions with other attributes other than hair friction, connect Sonification with technical data, and customize the Sonification approach for the target. Skincare and makeup categories are also in our scope of our next steps for future Sonification approaches.

We also plan to explore the potential of the Sonification in other countries in order to examine how sounds and music may convey similar or different perceptions and to extend this approach to other cosmetic products.

References

Brown D. et al (2011) Seeing with sound? Exploring different characteristics of a visual-to-auditory sensory substitution device. Perception, 2011, volume 40, pages 1120 - 1135

Court D. et al (2009) The consumer decision journey. McKinsey Quarterly

Kramer G. et al (1999) Sonification Report: Status of the Field and Research Agenda. National Science Foundation by members of the International Community for Auditory Display.

Kramer G. et al (2010) Sonification Report: Status of the Field and Research Agenda. DigitalCommons@University of Nebraska - Lincoln

Nomura M. et al (2016) Translating the human hair surface state into sound. IFSCC 2016 Congress

Song J (2010) Retail Design and Sensory Experience: Design Inquiry of Complex Reality. 2010 Design Research Society

Walus BP et al (2016) Sonification and music as support to the communication of alcohol-related health risks to young people. J Multimodal User Interfaces (2016) 10:235–246

The Authors

Sono Hasuka is Consumer Evaluation Researcher, Product Performance Evaluation, L'Oreal R&I Japan

Jung Yeon Nam is Product Performance Evaluation Manager, L'Oreal R&I Korea

Mariko Nomura is Instrumental Evaluation Researcher, Product Performance Evaluation, L'Oreal R&I Japan

Eri Yamashita is Qualitative Consumer Evaluation Manager, Product Performance Evaluation, L'Oreal R&I, Japan.

Damien Velleman is Instrumental Evaluation Manager, Product Performance Evaluation, L'Oreal R&I, Japan.

Video Storytelling in Medical Ethnographic Research

Uncovering unspeakable dynamics shaping patients' everyday lives

Afra Chen • Zou Xiong

Introduction

Patients have become the center of most pharma companies' worlds. Immersions are perfect for clients who want and need to understand the patients' world at the individual, customer level – to offer better cure solutions and better serve their patients with more customized products, this paper discusses the role of video-based research methods in medical research, why this method is particularly effective for helping the needs of those with multiple chronic conditions, to identify patient-centered improvement opportunities, and communicate them effectively to clinical and administrative leaders and decision making stakeholders.

Why visual anthropology matters

Until around the 1980s, mainstream anthropology was establishing itself as a scientific theoretical discipline. It had effectively become a mono-media anthropology, based on written text and verbal presentations. Within this written anthropology, however, by the 1980s some interesting debates had begun to emerge, with a developing interest in the body and phenomenology (in the work of Thomas Csordas), questions of experience (Turner and Bruner 1986), the senses (Stoller 1989), the status of text itself (Geertz 1988) and a continued insistence by applied and visual anthropologists of the value of their approaches. In these contexts, combined with the increasing availability and accessibility of visual media and technologies, visual methodologies and representations are more and more in use for both social science and applied science.

Figure 1. Still image from *Nanook of the North,* Robert Flaherty

The image from figure 1 is from the first ethnographic film Robert Flaherty made to give an unmediated presentation of native Alaskan culture. It is a daring move to feature how Flaherty worked closely with Nanook, his protagonist, to create a film based on their shared construction of Inuit everyday lives and dramatic events. The film is breath-taking back for that time. The audience is invited to experience how Nanook speared the seal, fought to get it and then eat the seal raw. It attracts so much interest and excitement that unfortunately can hardly be captured by written text.

Four key arguments stem from this groundbreaking approach:

1. Human's nonlinguistic forms of communication which typically involves some visual technology for data collecting and analysis

2. Some elements of human experience are best represented by visuals and which cannot be capitalized by text. The visual brings the fieldwork experience directly to the context of representation.

3. The final output, such as film, is a communicative activity and a datum of culture amenable to ethnographic analysis.

4. Visual media used for the presentation of data and research findings would otherwise remain verbally unrealized.

Ethnography and the patient experience

In the private sector, traditionally, marketing teams use insights gathered from market data and feedback from sales team and customer interactions to track market opportunities, customer segment and competition, etc., to create brand plans for increasing market access. In the medical sector, patient experience has often been long overshadowed in healthcare service delivery initiatives by the need to drive performance improvements and meet regulatory constraints. However, in recent decades China is undergoing vast changes and the whole infrastructure of healthcare is being reconstructed with changing demography, aging population, and increasing awareness of experiences other than medical treatment. This is encouraging medical practitioners/market stakeholders to reconsider their approach/relationship with patients. In recent years, the patient journey framework has been largely used in a growing number of projects with encouraging results in China. The sales & marketing team in the commoditized healthcare market faces the continuous challenge of increasing market access by differentiating their offerings. The typical patient journey depicts patient movement within and across healthcare providers by displaying interactions between patients and healthcare providers. It focuses on understanding how individual patients experience a disease and its treatment, and how patients learn about their disease, the kinds of decisions they make and, most importantly, the evolution of their emotional state as they move through treatment. Each time a patient interacts with healthcare services or products, the interaction is tracked and described. The goal is to collect required information only once, reducing the number of times the patient is interviewed, and increasing compliance to evidence-based best practices. Ethnography is known as a collaborative endeavor, and respondents are knowledgeable agents who both shape and are shaped by their communities and environments. They are the experts in the situation; ethnographers enter the field with open minds, wanting to learn from respondents, to record the process of reflexivity, collaboration and participation.

Visual empowered enthography research

However, during long years of practices, we found patient situations, especially for serious sufferers, are very hard to depict through a linear narration of text. Their complex condition and their daily fight nuanced with daily situations are too subtle for a business setting to learn enough about managing it. Emotions are also beyond text description most case scenarios.

In 2015, GSK approached Ipsos ECE for a study to understand COPD (chronic obstructive pulmonary disease) patient preference towards Seretide and Symbicort. Previously there had been no patient insight study conducted in the light of a holistic understanding of the patient's point of view; the physical & emotional motivations for their initial treatment & daily management, relationship dynamics (e.g. HCP & caretakers), decision making within their patient journey (i.e. prescription preference) and factors influencing patient adherence. We proposed the visual method as part of participant observation.

Most of our patients are elderly with low literacy to articulate complicated emotions. Some have high-esteem which keeps them from seeking emotional support from families and friends. Shrinking social circles after retirement keep them from a beneficial mutual relationship with social groups and activities.

Figure 2. Recruited patient for COPD patient journey study

Miao (74), Seretide/Ventolin user

Wang (59), Seretide user
Newly diagnosed

Jing (66), Seretide/Spirva user

Zhao (64)
Seretide/ Ventolin/Spirva user

Wu (69)
Seretide lasped user

Yuan
Seretide user

The challenges for us are that within a short period, we need to build rapport and allow them to convey their true feeling using their language and perspective. When patients show and tell about their suffering, and share anecdotes that evoke a strong response, it can hardly be reconstructed by a text-based report as the atmosphere, emotions, and context are all critical components for the effective conveyance of feelings. With video recording alongside the interview, we capture all these layers and bring them directly to stakeholders.

Griffiths suggests that Haddon's filmmaking was a form of 'haptic cinema' (through which the viewer 'feels' or 'touches' the image), which would have produced a 'sensorially rich' experience. (See figure 3.)

In the field, sometimes respondents find it hard to express what they are suffering and experiencing, and sometimes words just let the real atmosphere, context, and emotions slip away. We produce a 'flow of image of ritual events as they occurred', to let the 'intuitive assuming" cut through. Whereas with the help of the visual anthropologist, intuitive contexts are made visible and simultaneously recontextualise these moments of truth. The expression, gestures, and movements that naturally flow in the field represent a phenomenological and experiential holism that in many ways is often disrupted by a written ethnography. These are not staged events summarized by an analytic mind, but an unobstructed piece of moment that reflects individual nuances. Why do nuances matter?

Figure 3. Wang, 71, Seretide user, is showing and telling when he felt chest pain, how the massage bed helps to alleviate the level of pain

1. Within the same culture, differently positioned individuals might lead in-depth level of variations for context-based scenarios.

2. It can bypass general synthesis to reach deep seated essences of socially representative issues.

From the field we also learned that a lot of patients/ caretakers who are more experienced and informed think they can master the treatment without the advice of doctors. They are confident but do not always make the right decisions. However, this is a universal truth for most chronic disease sufferers, the conclusion is dry and open for deep understanding. We somehow lost the most essential layers of why they took the initiative to study about the disease and relevant medications through different channels, yet not believed in the doctors.

Figure 4. Chen (58) Wuhan COPD patient. Seretide lapsed user

Chen believes “胃食管返流“ gastroesophageal reflux is the root cause of all his problems, and take corresponding meds instead of COPD meds to deal with it. After years of experience managing COPD and navigating health systems in different hospitals, he thinks he knows pretty much all the drills. Doctors become the means to access the medications. But his perception and understanding of COPD is very likely wrong. By viewing through all his collected evidence, we still can’t be convinced by his wrong conviction, but after spending time with him we know that he has high self-esteem and is very independent, unwilling to rely on others for day-to-day life. He seldom talks about this physical status with a spouse, friends or peers. He lacks the desire to participate in social activities; he doesn’t want to become the burden of others. He is eager to find the ultimate “wonder drug”. In his deep mind he believes by tracing the root cause of COPD, it can be cured from its root. He visited different hospitals and different doctors since the diagnosis and has tried different treatments (TCM, health supplement, etc.). His life is somewhat centered around finding a good treatment of COPD.

He took us through his war trophies, different “learning materials” and notes and reports of exams he took. He thinks he knows the physical condition best and enthusiastically self-studies COPD from books, magazines and Internet, trying to be informed of the most pioneering and updated new treatments.

All these images convey his anxiety of losing control, facing suffering of sudden vulnerability. He just wants to regain control and become better. This enthusiasm and persistence is exhibited in his expressions, but behind his stubborn conviction is his deep un-spoken depression.

Figure 5. Self-study note of gastroesophageal reflux

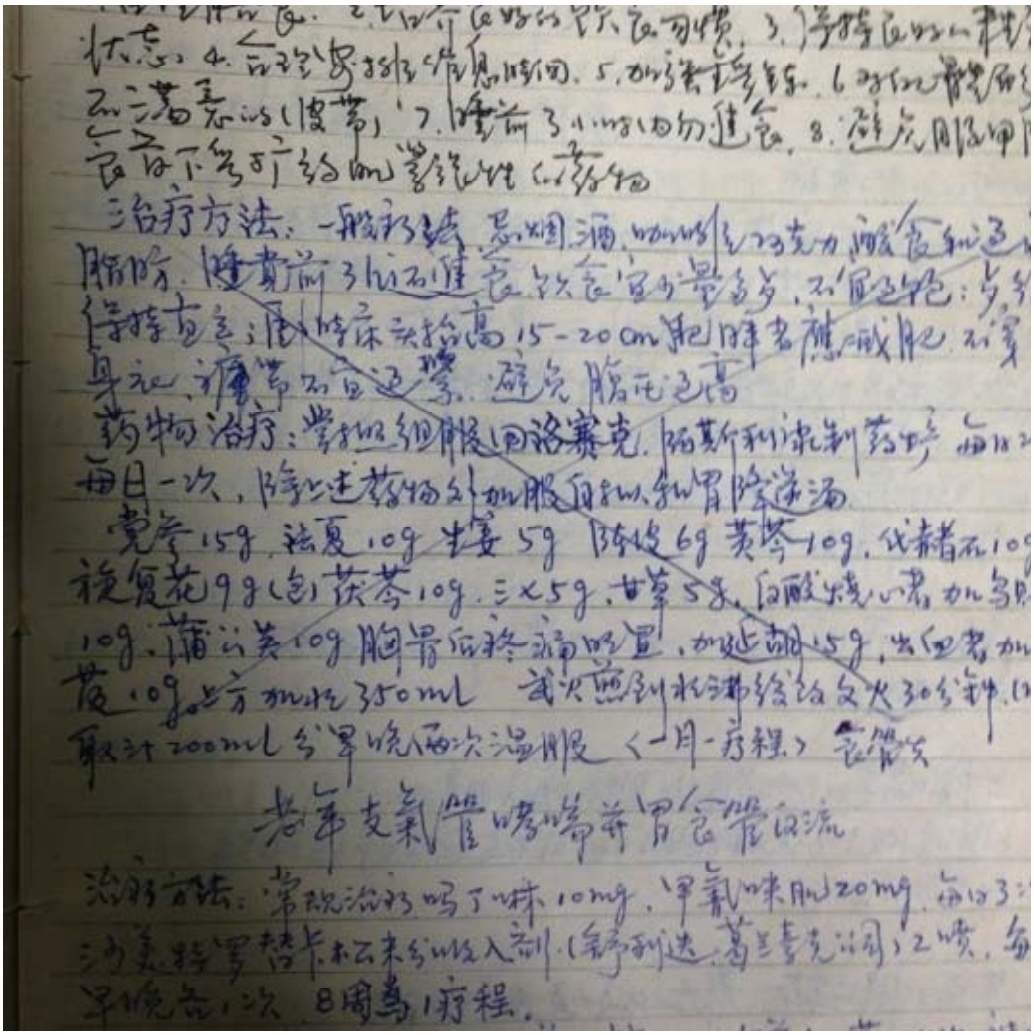
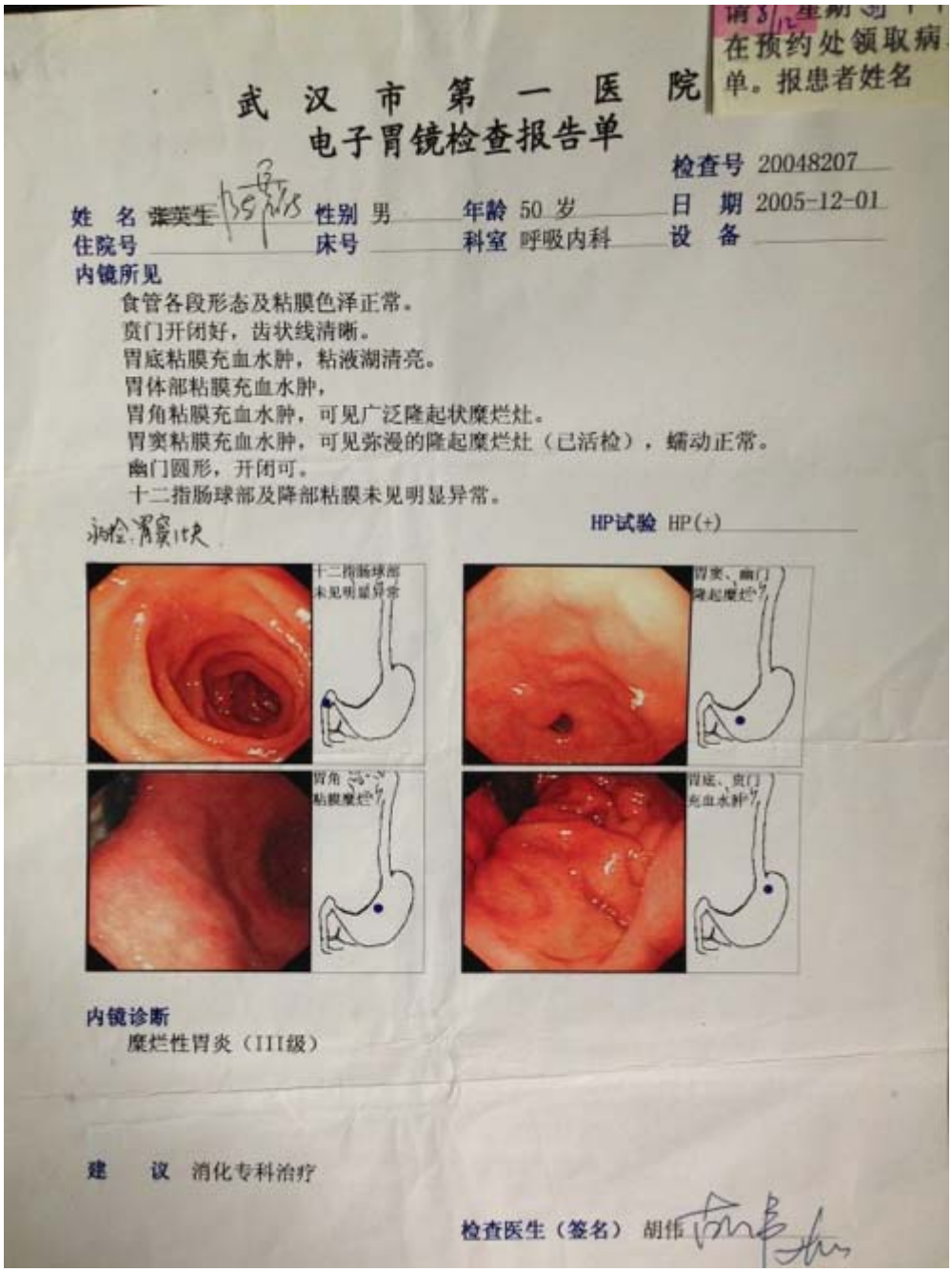


Figure 6. Test report of gastroesophageal reflux



武汉市第一医院
电子胃镜检查报告单

在预约处领取病单。报患者姓名

检查号 20048207
姓 名 张英生 性别 男 年龄 50 岁 日 期 2005-12-01
住院号 床号 科室 呼吸内科 设 备

内镜所见
食管各段形态及粘膜色泽正常。
贲门开闭好，齿状线清晰。
胃底粘膜充血水肿，粘液湖清亮。
胃体部粘膜充血水肿，
胃角粘膜充血水肿，可见广泛隆起状糜烂灶。
胃窦粘膜充血水肿，可见弥漫的隆起糜烂灶（已活检），蠕动正常。
幽门圆形，开闭可。
十二指肠球部及降部粘膜未见明显异常。

HP试验 HP(+)

内镜诊断
糜烂性胃炎（III级）

建 议 消化专科治疗

检查医生（签名） 胡伟

Without sharing the empathy of their stories, HCPs and manufacturers won't feel the sense of urgency of the need of patients to receive more explanation and consultation. They won't understand how the patient became disappointed and felt unsafe when the medical practitioner spent just a few minutes to diagnose and prescribe medicines. Manufacturers can compensate for this lack of trust in the medical system by providing more thorough/responsible consultation to assure long-term adherence, and bond with them emotionally and heartily. Ethnographic writing is mainly descriptive and reflexive and informed by theories and norms while ethnographic film is a quintessentially 'phenomenological medium' with a 'unique capacity to evoke human experience as reality' and not aw clears in the world of words. Sometimes we don't need any words to make an emotional decision.

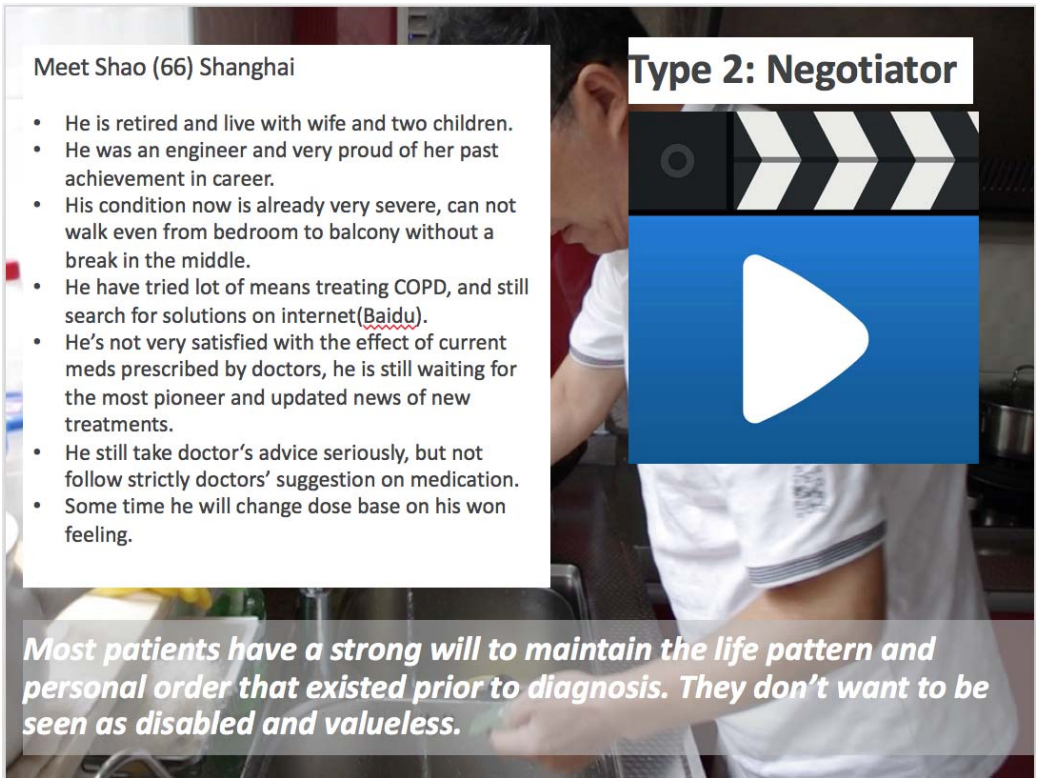
Another argument is based on the doubt that textual based media can fully capture patients' unspeakable suffering and subtle relationship with caretakers. Most patients have a strong will to maintain the life pattern and personal order that existed prior to diagnosis. They don't want to be seen as disabled and valueless. Confined by the illness, some patients are forced to step out of social circles, this becoming the most defeating experience for them. They try to guard their last little happiness even at cost of their health. This worries caretakers a lot and through the lens of the camera we see an invisible tension between beloved ones. To attempt to translate this into words would reduce its quality as knowledge and redefine the type of experience being represented. , Learning to help patients accept the illness and adapt to disease caused limitations are the first step to help the patient rebuild confidence and self-esteem, and to take a rational approach in seeking medical solutions. This not only alleviates the suffering of patients, but also the quality of life of their caretakers.

Figure 7. Zhao's (64) wife Jing (66), Seretide / Ventolin/Spirva user

Using visual storytelling goes beyond the static view of other text-based research models. It helps to visually bring in contextual insights around the patient experience, and promote empathy towards patient groups for those unarticulated emotions and tensions.

We use user personas and storyboards to feature the whole process of patient journey fieldwork. The final output is a 'short graphical description of a narrative' where a series of images are employed to represent a summary of patients' everyday life with diseases. (See figure 8.)

Figure 8. A typical illustration of key characteristics of user persona combing writing, sound, still and moving images



User personas involve creating caricatures of user groups in order to help stakeholders in better understanding the mentality of these groups, i.e. their 'expectations, prior experience and anticipated behavior'.

The storyboard distills research insights into a concise, visually compelling story of the customer's experience.

It is also helpful when working with hospital personnel to understand and improve the patient journey by representing information in a manner that is visual, interactive and comprehensible.

Conclusions

Our everyday life has become a "visual culture"', calling for a new field of study of both images themselves and 'the centrality of vision in everyday experience and the production of meaning'. The unique nature of patient experience also calls for a practice that is able to bring to the forefront ways of researching and representing other people's experiences and seeking ways to make these comprehensible to others by highlighting inconsistencies instead of generalization. Viewing recorded interviews and observations provides us with an opportunity to see nuances that in all likelihood we would miss otherwise.

Visual narratives facilitate a seamless immersive anticipation of this collaboration with its rich context and details of situation. However, for manufacturers, opportunities lie in providing good experiences to anchor their brand and product in the patient's heart. To differentiate their offerings, one needs to create a patient centric communication by understanding their target user with a more empathetic heart. Ethnography filming is a process of understanding achieved through a gradual process of discovery, that is, through engagement with the everyday lives of the subjects rather than by placing them within predetermined matrices. Understanding the context of patients and caregivers reveals discrepancies between what people say and what they do.

Patients and caregivers may voice confidence about managing medications or handling other care tasks, but their actions might reveal hesitation or misunderstanding. Viewing recorded interviews and observations as a team enables all stakeholders to compare their perceptions and generate a deeper understanding of how patients and families are really facing life challenges with real life scenarios.

References

Curry J , Fitzgerald A , Prodan A -《 Studies in Health Technology & Informatics》 – 2014 Combining patient journey modelling and visual multi-agent computer simulation: A framework to improving knowledge translation in a healthcare environment

(Johnston & Kong, 2011Johnston, R., & Kong, X. (2011). The customer experience: A road-map for improvement. *Managing Service Quality: An International Journal,*

Sarah Pink Routledge Taylor & Francis Group. The Future of Visual Anthropology Engaging the senses

The Author

Afra Chen is Head of Ethnography, Ipsos, China.

Zou Xiong is Senior Market Manager, GlaxoSmithKline, China.

The Hunt for an "Authentic" Coffee Experience

Using intuitive methods and a virtual experience technology to explore a total coffee experience

Junichi Ichiba • Humphrey Chen • Yuki Tanaka

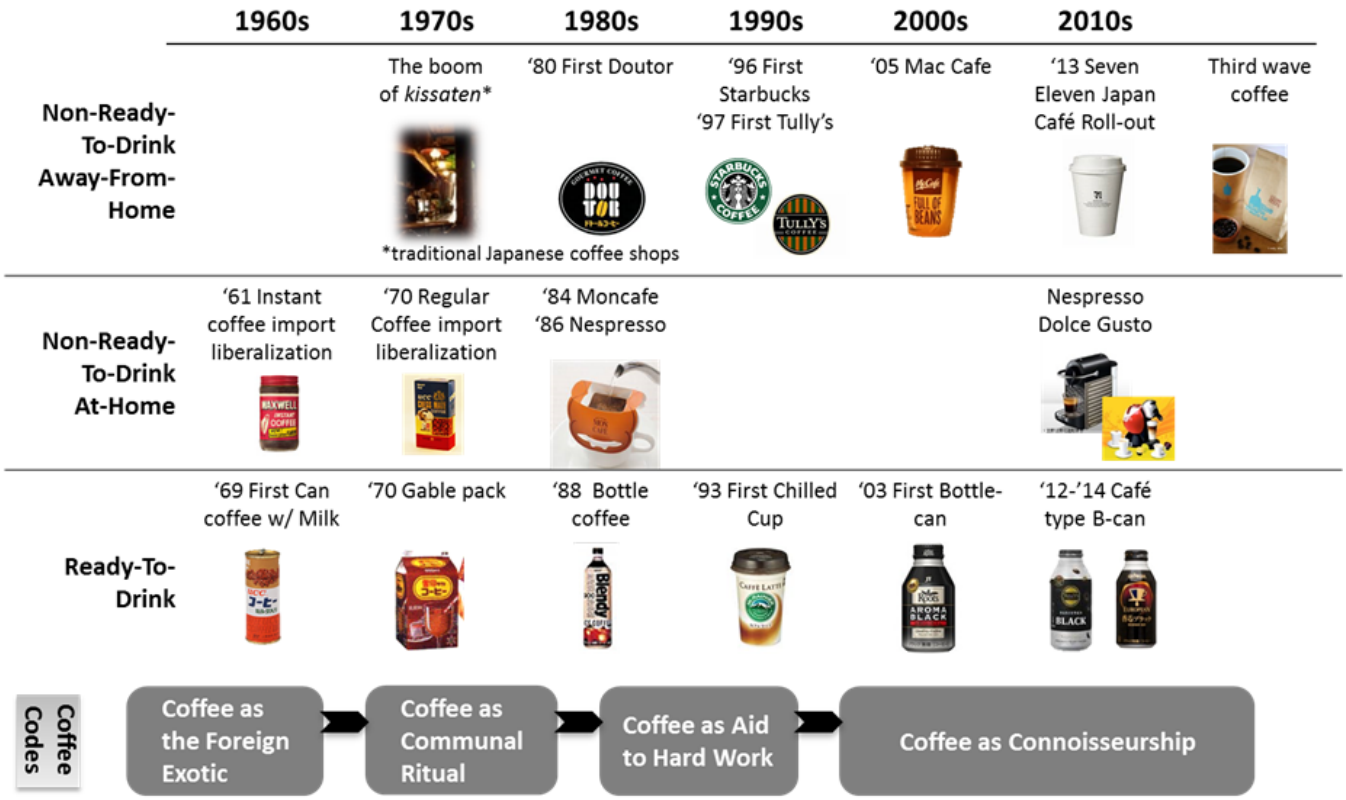
Introduction

Although consumers are quite particular when it comes to choosing coffee, understanding the criteria they base their choices on – particularly what causes them to consider coffee 'authentic' – can be difficult. Spurred by the convenience store industry's very successful move into the Japanese coffee market, Coca-Cola Japan (CCJC) saw offering its own 'authentic' coffee as an important growth opportunity. This paper describes how using innovative methodologies, CCJC/ INTAGE successfully gained insights into the above criteria and developed a new serving model.

The growth and evolution of the coffee market in Japan

The custom of drinking coffee started in the late 1700s in Japan, but it wasn't until the early 20th century when coffee shops first opened in large numbers, that the custom spread to the general population. For a long time, coffee remained something drunk in coffee shops. With the introduction of instant coffee in 1961, however, it became a household drink. Shortly afterwards, in 1969, the world's first canned coffee, "UCC Coffee", went on sale and immediately became a huge hit. With ready-to-drink (RTD) coffee, the beverage became something that can be drunk conveniently at any place, and the coffee market expanded greatly, driven by this convenience. As the market grew and evolved, the range of packaging also diversified to include carton packs, plastic cups, PET bottles, and "bottle cans" (metallic cans with a screw cap). (See figure 1.) Total coffee consumption volume increased steadily.

Figure 1. History of coffee in Japan



Vending machines (VM), ubiquitous throughout Japan, are an important sales channel for RTD coffee. At present, there are approximately 2.2 million RTD vending machines in Japan, offering not only RTD coffee but also Coke, non-sugar tea, energy drinks, and other beverages. These vending machines make it possible for people to enjoy RTD coffee conveniently at any time and any place.

Coffee dispensing vending machines first appeared in 1962 ─ before RTD vending machines ─ offering coffee and other soft drinks (see figure 2). The first CVMs dispensed liquefied instant coffee, but over time, machines that grind coffee beans and brew coffee were developed. There are now around 170,000 CVMs ─ roughly 8% of the figure for RTD vending machines ─ however, their numbers are declining.

Figure 2. A cup vending machine (CVM)

Subsequently, in 2011, convenience stores introduced “freshly-brewed coffee” (CVS coffee). (See figure 3.) When 7-Eleven Japan, the largest convenience store chain in Japan, implemented a nationwide rollout that brought CVS coffee to all of its stores (roughly 16,000 stores at the time), the other major convenience store chains quickly followed suit. By offering a convenient way for consumers to drink “authentic” coffee, CVS coffee spread rapidly, taking share from the RTD coffee market. CVS coffee is now available at nearly all of Japan’s 54,000 convenience stores nationwide, and has become firmly established as a part of coffee drinkers’ daily lives.

Figure 3. CVS coffee machine at Seven Eleven Japan

Business questions to reignite CVM for Coca-Cola system in Japan

The coffee market as a whole is growing, but the success of CVS coffee has resulted in a contraction in the RTD and CVM segments. For the Coca-Cola system in Japan, non-RTD coffee is important as a means of satisfying consumers' demand for "more authentic coffee" ─ a demand that cannot be met with RTD ─ and CVM is therefore critical as a channel for satisfying this demand. Developing a CVM that is competitive with CVS coffee is therefore a matter of great urgency for the Coca-Cola system.

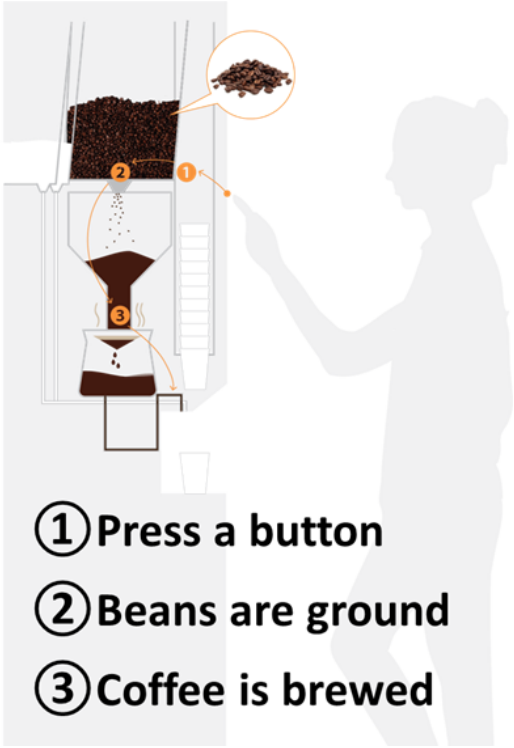
The key difference between CVS coffee and CVM coffee is presumed to be "authenticity," which is also a driver for growth for CVS coffee: the former is seen by consumers to be "authentic", while the latter is considered "inauthentic." However, in terms of the mechanics of how the coffee is made and served, CVM and CVS coffee machines are virtually identical; both grind coffee beans and brew coffee before dispensing it (see figure 4).

We therefore needed to answer two questions in order to set about developing a CVM that provides "authentic coffee" and is attractive to consumers and differentiated from other available offerings:

Question 1: What is 'authentic' coffee for consumers?

Question 2: Which prototypes would be felt to be authentic in the prototype testing stage?

Figure 4. Mechanical process of CVM and the coffee machines in convenience stores



Challenges for answering the business questions

Verbalizing 'the unconscious'

In carrying out this project, one of the biggest challenges we faced was identifying the criteria that people unconsciously use to determine whether or not coffee is "authentic", and being able to verbalize these criteria. Consumers may have some criteria that go into the choosing process to some extent, but choices are often made unconsciously, based on a vague feeling. When people are asked directly why they think something, they often say something like, "It's just a feeling" or "It's what I've always thought" in short, they themselves don't know why they think the way they do. Moreover, the fact that they say "It's just a feeling" suggests that the answer is not something that can be easily verbalized, such as the design or taste. Thus, conveying these underlying factors in clear terms was a major challenge.

Quality or efficiency?

The prototype testing stage presented another challenge: how could we test the prototypes in a cost- and time-efficient manner without sacrificing quality? Developing new sales equipment has generally been a process that involves the building of several prototypes and conducting a series of consumer tests. But it was soon apparent that building an actual prototype and testing it would be quite costly and time-consuming, and it would therefore be difficult to conduct a series of such tests. One commonly-used way to save time and money is to take the design alternatives that have been conceived in the prototype development process and conduct studies in which consumers are shown the designs utilizing the internet. However, such an approach has clear limitations, and in this case it was doubtful that a test using images much smaller than the actual object would elicit the same kinds of responses that the real object would. Thus, the challenge was to find an efficient research methodology that delivered a high level of accuracy at a minimum of time and cost.

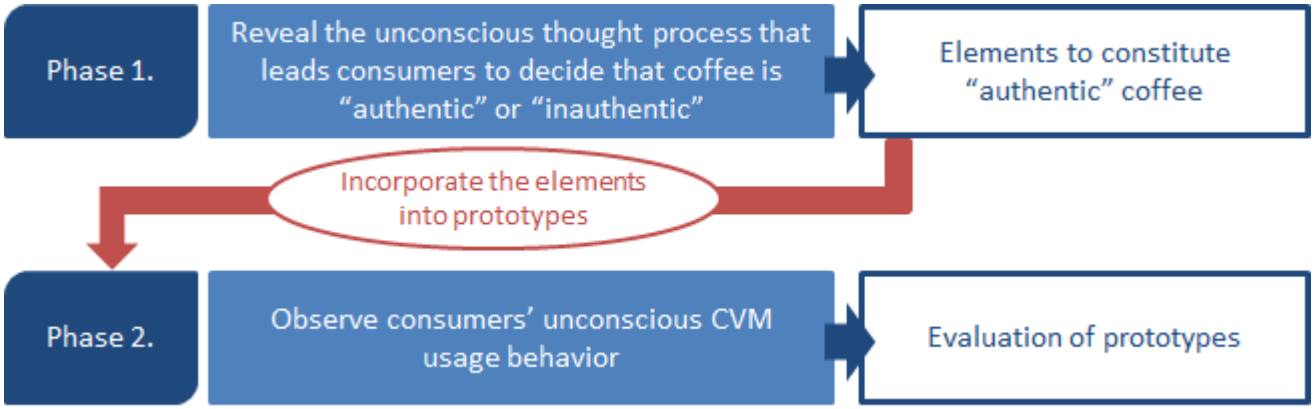
Our approach – Gaining access to 'the unconscious'

When posed a question, people try to find an answer. This is particularly true with research respondents, who attempt to provide the kind of answer they think is expected of them. For this project, we needed to

obtain two types of “unconscious output” from consumers using a methodology that did not make them feel they were being “questioned”.

Firstly, we needed to draw out the unconscious thought process that leads consumers to decide that coffee is “authentic” or “inauthentic”, and turn this unconscious process into something tangible. Then, our task was to have consumers go through the process of buying coffee from a CVM (without actually preparing large equipment), and observe their unconscious CVM usage behavior in order to judge the performance of the prototype alternatives (see figure 5).

Figure 5. Total research process



Facing these two challenges, INTAGE felt that it would be impossible to meet our client’s demands through traditional research methodologies, and therefore proposed the implementation of two innovative approaches.

Step 1: Revealing unconscious differences via the Sensitive Differentiation Method (SDM)

Step 1 is the phase to understand what consumers consider as “authentic coffee”, and thereby gain insight into what kinds of stimuli and experiences a CVM needs to offer in order to be perceived as something that delivers “authentic coffee.”

To be more precise, clear, explicit answers to the following questions are required:

- What is it about CVM that makes people see it as offering “inauthentic" coffee?
- What is it about CVS counter coffee that makes it “authentic”?
- To consumers, what is the meaning of “authentic” coffee?

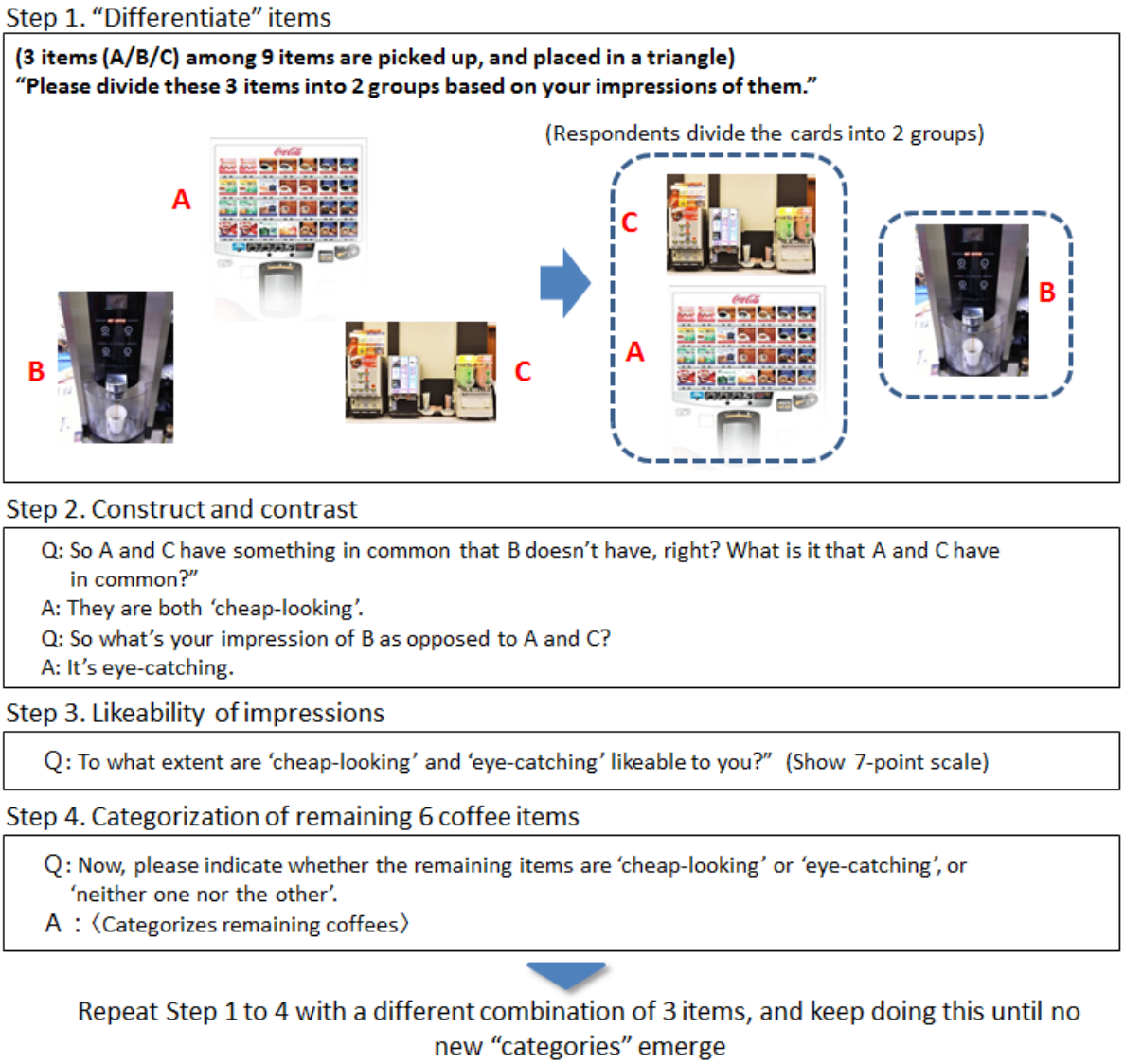
Previous and unsuccessful attempts indicated that direct questioning wouldn’t give us the answers. The approach we employed, therefore, allowed respondents to unconsciously define what is ‘authentic’ and what is not, according to their own criteria: Sensitive Differentiation Method (SDM).

SDM is an approach that uses a technique, originally developed for clinical psychology, to draw out subtle differences between brands and products that ─ since conventional analytical methods do not recognize them as different ─ are generally grouped into the same cluster.

To be more specific, it is an original methodology developed by INTAGE that employs the repertory grid technique. The repertory grid technique is a counseling method developed by George Kelly based on personal construct theory ─ i.e., the theory that people have cognitive 'constructs' that they use to understand things and make predictions about the future (Shaw, 1980; Rogers & Ryals, 2007). It is a method for eliciting a person’s 'constructs'.

The repertory grid technique involves presenting subjects with a series of three stimuli, or "elements". The subject compares the three elements and describes how they are similar or different. In this way, the subject's constructs are elicited through the subject's own words (see figure 6).

Figure 6. SDM process



The original methodology developed by INTAGE involved conducting one-on-one interviews as follows:

- Cards showing the major coffee products sold in Japan are prepared beforehand (nine cards for this project)
- The interviewer chooses three cards (coffee products) that consumers are unlikely to perceive as being different and shows them to the respondent.
- The respondent divides the three coffee cards into two groups (of two cards and one card).
- The respondent gives names to the two groups(for example, the respondent might name one group "Top sellers", and "Minor products")
- The respondent places the remaining six coffee cards into one of the two groups that were named as above.

With this methodology, consumers begin by making distinctions freely and spontaneously in their own minds, without being "questioned". There is no "verbalization" of the distinctions at this stage. Such distinctions are made apparent afterward, with the support of the interviewer, who ― without asking any questions ― helps the respondent see the criteria for differentiation that emerged in his or her own mind, and put this previously -unconscious criteria into words. This methodology therefore makes it possible to discover subtle yet important differences that are vague and difficult to describe.

As noted above, a wide variety of coffee products are consumed in Japan. For this study, we conducted interviews with 32 coffee drinkers and examined the following nine coffee product categories:

- CVM coffee: the main focus of this project
- Brewed coffee sold in convenience stores (CVS counter coffee): assumed to be the main competitor of CVM
- Starbucks coffee: the representative "café coffee"
- Kissaten or traditional Japanese coffee shop coffee: unique to Japan, and the model for Blue Bottle Coffee
- Fixed-price "drink dispenser" coffee provided at restaurants
- Home-brewed coffee
- Instant coffee drunk at home
- "Nespresso", the home espresso system from Nestle Japan
- RTD coffee

The coffee drinkers interviewed for this study were people who regularly drink at least three of these nine types of coffee, and enjoy drinking coffee of their own choosing at home and/or at work, whenever and wherever they like. Therefore, the respondents who spoke to us about the differences between the various coffee types were "coffee lovers" who are frequently choosing from among the various coffee options available according to some sort of reasoning process.

For this project, we also had respondents divide the constructs elicited via SDM into the categories "Like", "Dislike", and "Neutral" with the aim of incorporating the favorable elements of the product groups into the prototypes.

Step 2: Evaluation of prototypes via virtual experience

Step 2 is the phase to apply the findings of Step 1 to the development of several "authentic coffee" CVM prototypes, obtain consumers' evaluations of the prototypes, and select a prototype for advancement to actual product development stage, using the following criteria for selection:

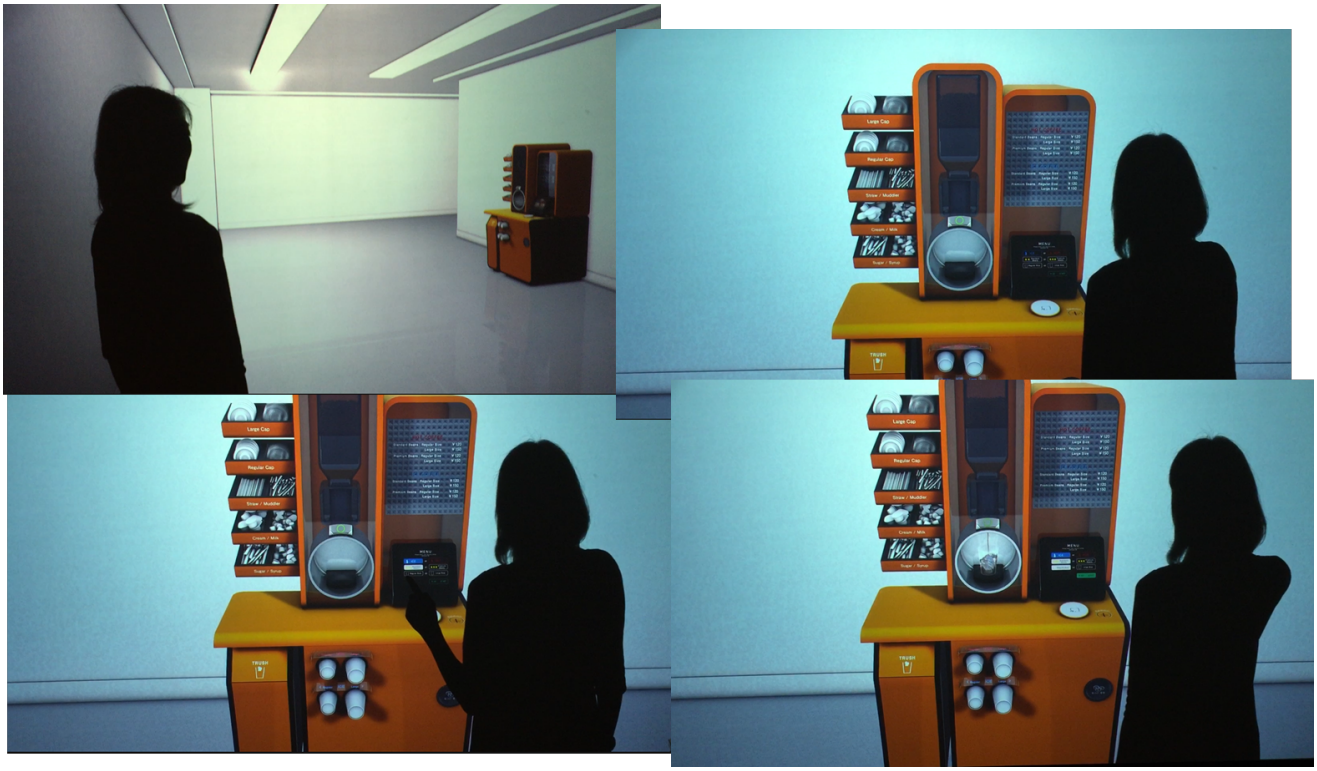
- Consumers have a more favorable impression of it, and want to purchase coffee from it.
- It is seen as providing authentic coffee.
- It is user-friendly and entices people to use it again.

As stated above, CCJC wanted a time- and cost-efficient approach that did not sacrifice quality. INTAGE responded to the client's demands by adopting a method that employs "virtual experience" technology and combines behavior observation with 1-on-1 interviews.

First, through the use of simulation technology that is capable of reproducing phenomena such as steam, liquid, smoke, flames, and flowing water, a prototype was projected onto a 120-inch monitor in high resolution (see figure 7). We then observed how the respondent acted in front of this new CVM that they encountered for the first time, without providing any explanation of how to use it. After respondents finished

performing a series of actions, we asked them to share their first impressions of the machine and their thoughts about its usability, as well as the reasons for and feelings behind any unusual behavior that they exhibited.

Figure 7. Virtual experience prototype: Code name “POP”



By seamlessly reproducing a realistic purchase experience ─ choosing a product, inserting a coin, listening to the sound of the beans being ground, etc. ─ we were able to observe consumers’ natural responses, such as surprise, confusion, etc., including changes in their facial expression. When respondents approached the CVM and pushed the product button, they could hear the sound of ice falling into the cup as well as the sound of the coffee being brewed, and as such they were able to experience the same kind of input that they would get from a test that used a real prototype, but without the extent of preparation that a real prototype would require. In this way, we were able to obtain responses that were close to real-life responses, and far superior to the kind of responses that would be gained had they interacted with a small image on a screen. In other words, the answers given by respondents after this virtual experience were reliable accounts of “actual experiences”. We were able to glean how they actually felt during this experience, feelings which emerged naturally from conversations they had with interviewers. Since this method costs only a fraction of what it costs to make a prototype, and the test period can be compressed by several months, we were able to test four different designs and brewing mechanisms simultaneously, which would have been impossible with real prototypes.

Moreover, the reduction in time and cost made it possible to recruit a larger sample. Altogether, we interviewed a total of 12 Japanese coffee drinkers with a wide range of tastes and backgrounds. We were able to present the prototypes to a wide range of coffee drinkers; to people who, for example, like canned coffee, are crazy about Starbucks, or like to make “authentic” coffee at home.

Discoveries through accessing “the unconscious”

As described above, in Step 1 we were able to identify the hitherto elusive criteria for making CVM more attractive and “authentic”. Moreover, since we succeeded in putting the factors that lie behind people’s vague feelings into words, it is now possible to sort this information and apply it to other activity as well.

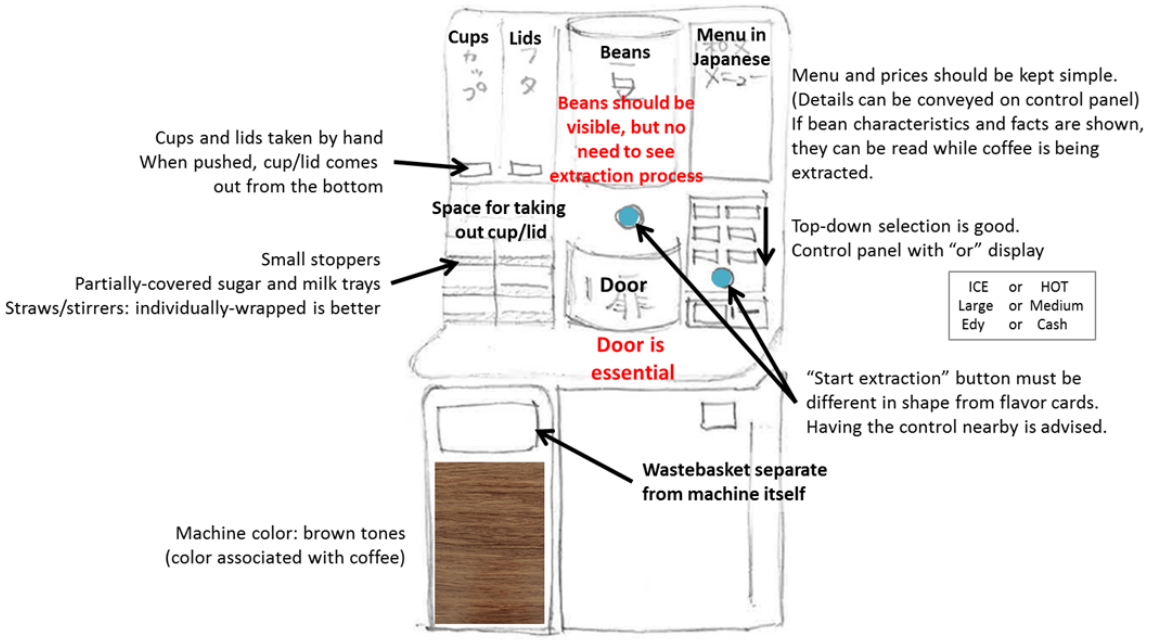
In Step 2, we succeeded in letting respondents experience the prototypes under development while compressing the cost and time investment. Through the re-creation of an environment that is similar to actual usage conditions, we were able to obtain realistic responses, so we could not only identify the essential criteria for creating an attractive CVM but also understand the feelings of consumers that these criteria are based upon.

In this study, we were able to discover what it is that makes a coffee feel “authentic” to consumers, and the series of experiences that strengthen this impression.

- For consumers, a large part of what makes a coffee “authentic” is the feeling that each cup is brewed individually, on the spot.
- Moreover, the feeling that each cup is brewed individually is based on a holistic experience that includes all five senses. The sight of the coffee beans (and seeing the number of beans decrease) and the coffee being injected, the sound of the coffee beans being ground and the liquid dripping down, the aroma of the coffee, the warmth of the cup when it is held ― all of this sensory input contributes to the impression that the coffee is being brewed, and is authentic.
- It is also important for users to form the impression that the CVM is a device for brewing individual cups of coffee, and not a machine that sells mass-produced items.
- Consumers’ perceptions of CVM are shaped by their past experiences with vending machines: when they see a CVM that is similar in shape and size to RTD vending machines, they unconsciously associate it with vending machines that sell mass-produced items. As a result, they get the impression that the coffee in CVM is mass-produced, rather than authentic.
- Participating in the coffee preparation process also contributes to the feeling that the user is “making" his or her own coffee, and therefore strengthens the feeling that each cup of coffee is brewed individually. Therefore, leaving some of the usage process to the user’s discretion rather than having the machine do everything automatically is more likely to produce a feeling that the coffee is being brewed individually, on the spot.

Since the above findings were expressed in clear-cut language and conveyed through visuals as well (see figure 8), the stakeholders in this project were able to come to a shared understanding. This in turn allowed for the creation of concept model ideas and the test marketing of a real prototype.

Figure 8. The CVM image, based on the findings from Step 1



In addition, in order to speed up the roll-out process, a “See-through CVM” —a new version of the current CVM that incorporates the findings of this project—has been developed, and is also being deployed (see figure 9).

Answers to the business questions

Did the approach answer to the business questions raised by CCJC? It is fair to say the answer is “yes.”

Verbalizing ‘the unconscious’

The criteria upon which people make unconscious decisions were successfully discovered, as well as the principles of what makes coffee authentic. The discoveries enabled CCJC to develop prototypes that bring these findings into life.

Quality or efficiency?

The virtual experience technique dramatically reduced the cost and time to test the prototypes – CCJC tested them for about one-third of the cost and time of producing actual prototypes. Being cost- and time-efficient, it enabled them to test more numbers of prototypes than usual in a way that let people feel they were actually using a machine.

Concluding remarks

As an application to other research, the SDM method is applicable to clarifying and articulating subtle differences. For example, a package perception research among RTD coffee in Japan such as can, Bottle-can, PET or carton packs. The virtual experience method can be applied to research which requires reproducing real situations like as SME or POP developments. This method can gain more natural and intuitive reactions from respondents.

Figure 9. “See-through CVM”

References

Rogers, B., & Ryals, L. (2007). Using Repertory Grid to access the underlying realities in key account relationships. *International Journal of Market Research, 49(5),* 595-612.

Shaw, M. L. (1980). *On becoming a personal scientist: interactive computer elicitation of personal models of the world*. London: Academic Press.

The Authors

Junichi Ichiba is a Senior Manager, Coca-Cola (Japan) Company, LTD.

Humphery Chen is a Sales & Innovation Director, INTAGE Inc., Japan.

Yuki Tanaka is a Sales Analyst/Coca-Cola (Japan) Company Account Manager, INTAGE Inc., Japan.

PART 3
Discovering the Undiscovered!

The Missing Link

A breakthrough to extracting emotional drivers of behaviour

Brian Fine • Con Menictas

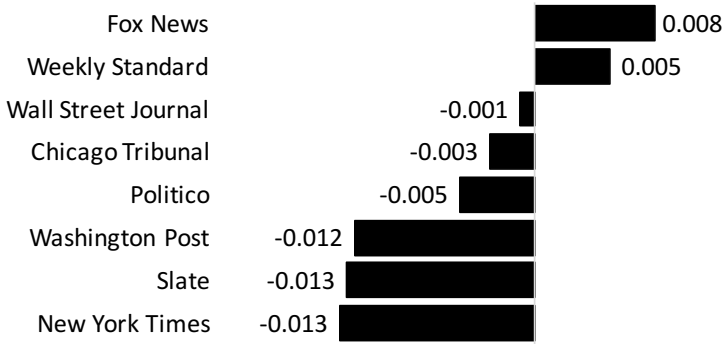
Introduction.

The 2016 US presidential candidates

Emotions are the foundation of what makes us human. Emotions determine how we think and behave. The emotions we experience each day drive us to actions. These actions can be important or trivial, and everything in between (Myers, 2004).

Recently in the US elections, we saw how the media tried to play on people's emotions. Irrespective of what Trump or Clinton did, the media generally portrayed Trump in a negative light and Clinton in a positive light. To illustrate the point, figure 1 shows the positive versus negative sentiment indexes towards Trump portrayed by the media over the entire US election campaign (Sides, 2016)

Figure 1. Indexed campaign media sentiment towards Trump



According to both candidate camps, however, each claimed that the other's attendances were inflated by the media. The national polls closest to the election, including those conducted by ABC News/*Washington Post*, Ipsos, YouGov and Fox News, all estimated a Clinton lead of 3–4% over Trump (Skibba, 2016).

What clearly emerged as the deciding population proportion that got Trump over the line were the swing states in rural areas and small towns. These were the most hard-hit financially, the worlds where dreams are forever lost and emotions run high – the forgotten states that Washington seemed to ignore, i.e. the states of Michigan, Wisconsin and Pennsylvania (Hanrahan, 2016).

Judging the speaking techniques that both candidates used, Clinton spoke as the professional and seasoned Congressperson that she is. Clinton engaged polished, logical and rational appeals, such as "*If you believe that we should say "no" to unfair trade deals ... that we should stand up to China ... that we should support our steelworkers and autoworkers and homegrown manufacturers … join us*."

Trump, on the other hand, used a completely different approach - much simpler, easier to understand and immediate and full of non-sequiturs, thus making each statement stand on its own. Trump's sentences often had no ending, letting his audience finish them for him – all this making people understand Trump on an emotional level. As Mark Liberman, the famous linguist from the University of Pennsylvania (Golshan, 2017), said *"No need for sacrifice, for compromise, for complexity. He taps into fear and insecurity, but then enables his audience to express that fear through anger. And anger gives the illusion of empowerment."*(Mez Du, 2016).

Our capacity as humans for emotions has evolved through natural selection, due to their value for survival and reproduction (Darwin, 1875; Ekman & Friesen, 1975, 1982). Fundamentally, there are six basic emotions: (1) Surprise; (2) Fear; (3) Disgust; (4) Anger; (5) Happiness; and (6) Sadness (Ekman & Friesen, 1982).

Notice how Trump taps into these six basic emotions with very simple sentences. For instance:

1. *"This is very bad"*;
2. *"We've made other countries rich"*;
3. "*They're bringing drugs"*;
4. *"They bring crime"*; and
5. *"America's working class is dying"* (SMH, 2017).

Contrast Trump's simplicity to Clinton's complexity:

1. *"Making our economy work for everyone, not just those at the top, protecting our country and protecting our planet"*;
2. *"We spent a year and a half bringing together millions of people from every corner of our country to say with one voice that we believe that the American dream is big enough for everyone"*; and
3. *"Our Founders fought a revolution and wrote a Constitution so America would never be a nation where one person had all the power."* (NYT, 2016).

Of course, we have chosen snippets from the extremes of each candidate's speeches to illustrate our point. However, notice that apart from the very short sentences that Trump uses, his language is extremely emotive. It engenders fear, surprise and anger.

Clinton's longer and multifaceted sentences do not immediately tap into the six basic emotions mentioned earlier. They are grandiose statements of fact - extremely important of course - and even though all-encompassing and inherently presidential, seem more like stories that write themselves. They are more akin to lighting the first fire that warms the room, rather than blinding strikes of lightning that violently shake a mountain's foundation.

Many suffering Americans have said of Clinton that her statements might be best left in story books, rather than them being cries for change and fairness that the common trampled worker has intensely longed for. They do not get under the skin of the electorate like Trump's molten one liners do. They do not say the things that the forgotten people of America want said.

The frustration and anti-institutional feelings that drove the Trump campaign may also have aligned with an unwillingness to respond to polls (Mercer, Deane, & McGreeney, 2016). Some have also suggested that many of those who were polled simply were not honest about whom they intended to vote for. The premise of so-called *shy-Trumpers* suggests that support for Trump was socially undesirable, and that his supporters were unwilling to admit their support to pollsters.

This hypothesis is reminiscent of the supposed *Bradley effect* (G. Payne, 1986), when Democrat Tom Bradley, the black mayor of Los Angeles, lost the 1982 California gubernatorial election to Republican George Deukmejian. This was the case despite having been ahead in the polls, supposedly because voters were reluctant to tell interviewers that they were not going to vote for a black candidate.

In Trump's case, standard polls during the campaigns seem to have failed to pick up Trump's lead, because they followed a traditional and less emotive approach. We propose that an approach which accounts for constituents' emotions and their emotive response to each candidate could have worked a lot better in picking up the true sentiment for Trump over Clinton.

Emotions played a big role in the 2016 US election. Dr Cosgrove of Suffolk University in Boston said that the 2016 US presidential election proved the power of emotional branding, positioning and understanding the strategic conditions in which a campaign is run (Cosgrove, 2016).

We present findings from both a qualitative and quantitative study. In the qualitative study, we conducted online qualitative groups followed by a quantitative questionnaire administered to informants. We did this to gain a rank order of importance of the qualitative outputs. Following the qualitative study, we conducted a quantitative study of free form open text followed by suggested sentences, administered to respondents as an opportunity to verify their text inputs.

We propose that the suggested sentences offered an enhanced approach to gaining insights from open text when offered to respondents immediately following their text input. In cases where there is a strong emotional content in what is being researched, we demonstrate that emotions are captured using free form text just as well as they are in qualitative research, because open text is an unrestricted form of preference elicitation. In addition, open text offers the opportunity to be able to statistically determine the significance of themes, due to large samples, which is usually indeterminable in the low samples of qualitative research.

Against this backdrop, let us next examine how we have attempted to extract meaning and emotion from the open text data that we collected.

Extracting meaning and emotion from text

The formation of emotional factors, or latent variables of emotion that one uses to predict customer behavior, is ordinarily based on a set of questions derived from exploratory work. This exploratory work is by and large qualitative in nature. It is based on focus groups or in-depth interviews, where informants are probed to extract meaning concerning a given topic.

More recently though, open text data has been used to supplement, or in certain cases replace, the qualitative research function altogether. However, unless open text is processed by way of humans coding the text data itself, we are at the mercy of text software to do the job.

However, this has limitations. Even though text software is great for producing frequency data, as in the case of Wordclouds (Halvey & Keane, 2007) and in some cases related themes, rarely is it ever used to form and crystallize unique variables beyond isolated words, which is not much use for the formation of factors or latent variables.

On the other hand, the alternative approach of text coding by humans is subjectively biased to the coder's assessment of how a widely disparate set of responses may be grouped. Human coders, however, exhibit low accuracy because no two coders produce the same results. Humans have memory limits and these can be severely affected by the individual coder in question. On the positive side, when humans are performing optimally, and have perfect understanding of complex meaning, it is great for text mining, but how can we be sure that they are performing? This therefore calls to question the effectiveness of human coders. Moreover, coding takes place after completion of the quantitative survey, at which point no further information may be garnered from respondents.

Machines, on the other hand, meaning text analytics and mining software, have high accuracy because the same coding heuristic rules apply always. Machines have the capability for an unlimited number of codes. However, they have a low level of understanding of complex meaning ("Odintext Inc.," 2016).

We therefore propose a new approach which incorporates codification of open text on-the-fly of the spontaneous responses during the fieldwork. This new approach captures fresh factors coming out of early open text responses and subsequently presents them as text response suggestions to subsequent respondents as they are completing their text response.

In executing this process, we provide respondents the option to select from a range of suggested sentences as options best representing their text response as soon as they have completed entering their open free form text. This contradicts the post-field interpretation of open text responses, which is what happens when a human coder attempts to group disparate text after the study has been completed.

Suggestions are offered to the respondent as a supplement to the open text response they have provided, to both aid the respondent to confirm their most immediate response, and to ensure that the breadth of insight retrievable through open text is captured more fully, as shown in table 1.

Table 1. Process of raw data conversion into sound

Words	*Suggested Sentence*
Honesty	This candidate seems to be sincere
Integrity	

The aim of our approach is to capture, in a more reliable way, the key factors that drive customer behavior via suggestions offered to the respondent in real time. This approach has the effect of forming factors or latent variables on-the-fly.

We base our work for this paper on the most recent US election. The results suggest that in addition to a wider set of measures, emotional factors can be extracted from open text which aid in the execution and analysis of a quantitative survey, by using *enhanced* i.e., real time open text suggestions during open text data collection. Our work demonstrates that improved insight into emotional drivers of behavior is achievable through this enhanced approach, when compared to alternate methods of open text analysis and qualitative techniques.

Having introduced our topic and discussed the potential enhanced open text data extraction, let us next examine the data collection method, and subsequently the approach we undertook for the analysis.

Design

Sample frame

We collected data across all 50 states of North America at two time periods, using both quantitative and qualitative samples.

The total quantitative sample amounted to N=9,980, meaning roughly n=100 per state per period, as illustrated in table 2.

Ten online focus groups were also carried out, resulting in a total of N=100 informants. One online focus group for each of the 10 randomly selected states of the 50 US states, over two time periods, as illustrated in appendix 1, were executed.

In both the quantitative and qualitative sample design, we attempted to cover the broadest spectrum of North America, because we understood that many factors would come into play in terms of understanding how people felt, and that these would vary geographically. We recognized that the volatility of public sentiment was being fuelled by subjective media reporting resulting in media bias (Eberl, Boomgaarden, & Wagner, 2005), the volatility of the rust belt (High, 2003) and especially the aggressive attack ads employed by both candidates in specific US states (Geer, 2006; McGuffen, 2004).

Table 2. Quantitative sample sizes across the 50 US states

State Count	State	Time Period 1 Sample Size	Time Period 2 Sample Size	State Count	State	Time Period 1 Sample Size	Time Period 2 Sample Size
1	Alabama	95	98	26	Montana	99	101
2	Alaska	98	98	27	Nebraska	100	101
3	Arizona	100	98	28	Nevada	103	104
4	Arkansas	99	104	29	New Hampshire	97	99
5	California	104	100	30	New Jersey	100	102
6	Colorado	104	101	31	New Mexico	102	96
7	Connecticut	98	101	32	New York	96	95
8	Delaware	100	95	33	North Carolina	99	95
9	Florida	98	99	34	North Dakota	96	102
10	Georgia	104	98	35	Ohio	102	97
11	Hawaii	98	104	36	Oklahoma	103	99
12	Idaho	101	95	37	Oregon	104	101
13	Illinois	102	101	38	Pennsylvania	100	104
14	Indiana	101	97	39	Rhode Island	102	104
15	Iowa	103	103	40	South Carolina	102	101
16	Kansas	99	103	41	South Dakota	96	104
17	Kentucky	103	104	42	Tennessee	95	97
18	Louisiana	101	103	43	Texas	103	95
19	Maine	102	98	44	Utah	95	100
20	Maryland	101	95	45	Vermont	96	99
21	Massachusetts	95	99	46	Virginia	104	100
22	Michigan	100	98	47	Washington	97	100
23	Minnesota	98	104	48	West Virginia	100	103
24	Mississippi	103	102	49	Wisconsin	98	104
25	Missouri	95	96	50	Wyoming	97	95

As a result, we opted for a systematic spatial random sampling strategy (Stein & Ettema, 2003), rather than concentrated samples in main areas of population densities, as evidenced in table 2 for the quantitative study and appendix 1 for the qualitative study.

Research design

The qualitative component comprised two major parts.

First, the qualitative group discussions, where informants were investigated on their opinions about Trump and Clinton. Qualitative approaches were used such as word associations; sentence completions; constructions and expressive techniques (Malhotra, 2010), to allow informants to more fully express the way they felt towards both candidates.

Second, the focus group informants also answered Likert-type monadic rating scale questions that were developed to contrast both candidates, Trump and Clinton. Three analysts were used to codify the qualitative outputs and convert them to Likert-type statements representing the qualitative outputs. Each analyst prepared their own list of statements and then the statements were cross examined to establish a unified expression of statements, thus avoiding subjective bias during the formation of the statements.

We needed to ensure that there was sufficient agreement amongst informants of the sentiment towards the two electoral candidates. Also, we wanted to be able to make a direct comparison between the qualitative outputs (even though in quantitative form), to the emotional suggested factors or latent variables that would emerge from the quantitative free form text inputs. We therefore administered the Likert-type questions to the informants so that we could establish an order of importance of the qualitative outputs.

The quantitative component comprised two parts also.

First, respondents were asked to provide open text responses for both Trump and Clinton on the question: *"Why would you vote or not vote for Trump?"* and then the same question again for Clinton. The question we posed was presented in the same form of a classical open text approach. As an example, in a typical NPS survey (Reichheld, 2003) we might ask the question: *"Would you recommend this brand to your colleagues, family and friends?"*, on a Likert-type monadic rating scale such as: {1 = strongly disagree, 2, 3, 4, 5, 6, 7 = strongly agree} (Likert, 1932). This would immediately be followed by an open text question asking: *"Why did you give this rating?"* So with our question. Respondents were randomly assigned to be asked about Trump or Clinton first, followed by the other candidate in order to account for order effects (S. M. Friedman, 2012).

Second, as the initial respondents completed their open text input, we collated the most common responses and offered them as suggested sentences, on-the-fly. The enhanced text data therefore were generated from suggestions that were offered to the respondent in real time during the text data collection. As with the qualitative Likert-type scale formations constructed by three analysts discussed previously, so with the quantitative session's suggested sentences. Three analysts were used to individually construct the sentences and then a unified set of sentences were agreed upon, on order to avoid the subjective bias of each analyst acting alone. In order to make the codification of nearly 10,000 open text responses operationally and economically feasible, we used a human assisted computer coded text analyzer, meaning that a text analytics engine processed the open text data into meaningful themes and then the analysts put them into sentences.

This is in stark contrast to a typical NPS approach, where only an unbridled request for open free form text responses is sought. In our case, we offered the facility of open free form text to be entered, followed by suggested sentences that best represented the open free form text.

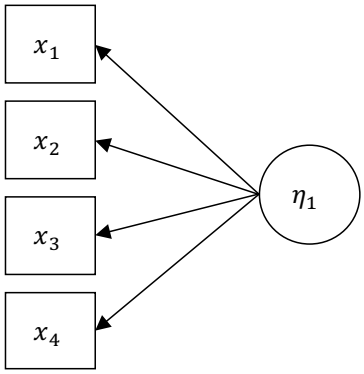
For example, where a respondent provides an open free form text response that might indicate that Trump is forceful or wants his own way, a subsequent suggested statement might be made along the lines of *'Trump is more aggressive that Clinton'*. As the respondent sees this suggestion, having completed their open text response, the respondent might think that the suggestion closely encapsulates their thinking and hence selects it also.

In so doing, we increase the frequency of a smaller set of terms, which does two things. First, we achieve an enhanced degree of agreement amongst respondents. Second, we get closer to forming factors of sorts, i.e., we tap into the underlying latent variable that drives the response.

A factor or latent and unobserved variable η, and also known as the true score x_t towards an object being evaluated, drives the rating on the observed score x_o (Hambleton, Swaminathan, & Rogers, 1985). Similarly, a suggested response immediately following an open text response, attempts to mimic a latent variable that underlies the open text response, thereby emulating a factor or latent variable. Figure 2 illustrates this graphically.

We see that η in fact is the underlying driving force or the latent or factor variable (Tinseley H. A., 2000) of the x_O variables i.e., the variables that contain the raw respondent quantities collected from respondents. So with text data, the literal responses are the x_1 through x_4 variables and η_1 is the factor or latent variable.

Figure 2. Example of a latent or factor driving observed scores



Therefore, when we present a proxy for the factor or latent variable to respondents by way of a recommended sentence which captures the meaning of what respondents are trying to say, we have essentially constructed a factor or latent variable.

Dynamic coding process

The dynamic coding process involves two steps.

First, the machine coding begins with the tokenization of individual words, which is the process of breaking a stream of text up into words, phrases, or other meaningful elements called tokens. The list of tokens becomes input for further processing.

Following tokenization, stop words are removed which are largely prepositions and pronouns using a variety of methods such as Zipf's Law, mutual information and term-based random sampling (Vijayarani, Ilamathi, & Nithya, 2015). When the stop words are removed, the remaining words are stemmed, meaning the root or stem of the word is revealed. As an example, the word *'catty'* would be stemmed to the word *'cat'*.

Second, once the words are prepared, a hierarchical cluster analysis is conducted to group the words into groups. These groups are then presented to the analysts which convert the groups into every-day language. To illustrate the point with an example, in table 1 earlier, the words 'honesty' and 'integrity' were clustered together, meaning that according to the body of text, they were related. The individual analysts, by agreeing between themselves as to the ideal representation of the two words in every-day-speak, converted the two words into the sentence 'This candidate seems to be sincere'.

The text analysis approach that we are using does not require an exhaustive text mining and analysis procedure. Humans take over from the machine at the point where the machine has produced themes or categories. This intervention by humans provides a much faster and more accurate representation of what the respondent is trying to say. The speed gain is transferred to an almost instant presentation of suggested responses to the next respondent. The full extent of suggested responses is quickly established, usually within the first 50 respondents at most, which for an online survey is effective in both cost and time.

Analysis

Our analysis examines and compares four different sources of text collection which may be used to find meaning and emotion driving consumer behavior. The first three of these follow traditional and modern methods, while the fourth is our proposed enhanced approach.

1. Qualitative focus groups – content analysis (traditional)
2. Quantitative rating scale – key outputs (traditional)
3. Quantitative open text response – standard Word Count method (modern)
4. Quantitative enhanced data via suggested sentences (enhanced)

Let us discuss each of these in turn:

Qualitative focus groups

To set the stage, we first summarize the campaign policies for each candidate as illustrated in table 3, where items 1 through 6 were the *most* mentioned for each candidate during their campaigns.

Table 3. Prelection policies for Trump and Clinton

Item	Trump	Clinton
1	*The wall*	*Boost the economy*
2	*Temporary ban on Muslims*	*Strengthen international relations*
3	*Deport illegals*	*Soften immigration policy (keep families together)*
4	*Bring American jobs back*	*Criminal justice reform*
5	*War on ISIS*	*Mental health support*
6	*Cut taxes*	*Domestic terrorism*
7	Health care (repeal Obamacare)	Student loans
8	Global warming	LGBT rights
9	Renegotiate trade agreements	Gun violence
10	Import tariffs for China and Mexico	Wall street reform
11	Harder torture	Income inequality
12	LGBT	Gay marriage (originally only civil unions)

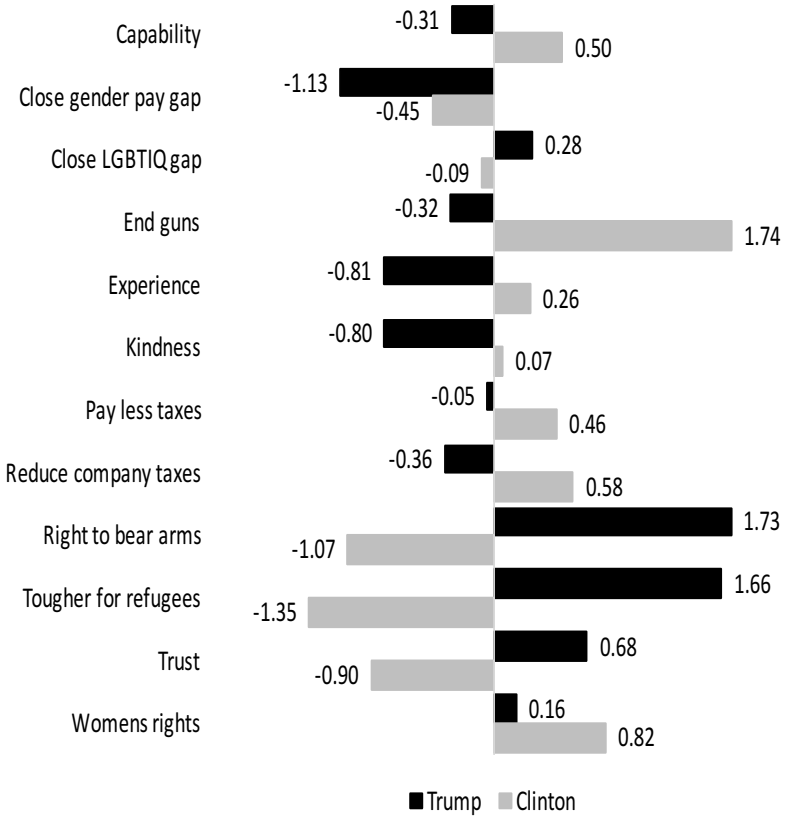
A review of comments made in the qualitative focus groups reveals that across the 50 US states, the things that separated the candidates were a lot less than the 12 most cited policies for each candidate on their respective US election campaign websites. This is illustrated in figure 3, which shows the qualitative output in the form of content analysis (Krippendorff, 2004). We see that (1) Capability; (2) Experience; (3) Trust; (4) Taxation; and (5) Kindness or Empathy, were the features that respondents were most concerned about i.e., the highest frequencies, when thinking of how to vote for the candidates. As we will discuss soon in the quantitative outputs, our enhanced approach also reveals factors that tap into the personal character of the candidates, more so than quantitative word counts alone.

Quantitative rating scales used in qualitative sessions

As we discussed earlier, the qualitative group participants were also administered a quantitative questionnaire to retrieve the order of importance of the qualitative outputs, the results of which are shown in figure 4. We mentioned previously that three analysts compiled the Liker-type rating scales on-the-fly individually, followed by a unified agreement on the final scales that were presented to informants in order to avoid bias.

Note how the qualitative content analysis in figure 3 has the highest frequencies on the emotional elements e.g., the feeling the candidate is capable, or has the experience, or can be trusted and whether they are kind. Contrast these with the *complete* set of elements that arose from the qualitative session's rating scales in figure 4. Notice how when asked about the same elements in a quantitative way i.e., the rating scales, the emotional elements lose their prominence. This suggests that the heuristics informants and respondents employ vary in a qualitative setting when compared to a quantitative survey respectively (Goldstein, 1990; J. W. Payne, Bettman, & Johnson, 2002; Thurstone, 1927, 1928).

Figure 4. Likert-type rating scale statements presented to focus groups (indexed)



Quantitative word counts

We next consider the word frequencies modeled by ordinary least squares (OLS) regression for both candidates from the quantitative study's open free form text. In Clinton's *and* Trump's case, the following words occurred in high frequencies as shown in table 5, with 33 words from the respondent response themes, ranging from (1) Aggressiveness through to (33) Overall Preference.

Looking at the significant words only i.e., t-values of greater than the absolute value of 1.96, we reduce table 4's word listing to what is important for each candidate in table 5. Note that we have a fair degree of similarity between the qualitative outputs and significant words. For instance, (i) capability; (ii) experience; (iii) trust; and (iv) taxation.

Table 4. Individual words used in OLS regression for both candidates (Sig. >1.96 t-val.)

Item	Words	*Trump*		*Clinton*	
		Coef.	t-val.	Coef.	t-val.
1	Aggressiveness	1.34	1.58	n.a.	n.a.
2	Arrogance	-0.65	-1.06	n.a.	n.a.
3	Capability	n.a.	n.a.	0.96	2.98
4	Competence	0.54	1.62	n.a.	n.a.
5	Compromises	-0.16	-0.26	n.a.	n.a.
6	Democracy	0.98	1.18	n.a.	n.a.
7	Ego	0.49	0.90	n.a.	n.a.
8	Email Scandal	n.a.	n.a.	1.18	1.24
9	Establishment	n.a.	n.a.	1.59	1.77
10	Experience	0.63	3.12	1.03	5.25
11	Guns	2.65	2.20	n.a.	n.a.
12	Honesty	n.a.	n.a.	0.59	0.91
13	Integrity	n.a.	n.a.	-0.89	-4.23
14	Intelligent	n.a.	n.a.	0.34	0.46
15	Leadership	0.93	3.77	0.38	0.78
16	LGBT	0.16	0.25	n.a.	n.a.
17	Likeability	1.02	5.36	1.19	5.40
18	Minorities	2.65	2.20	n.a.	n.a.
19	Morals	0.93	1.31	n.a.	n.a.
20	Muslims	0.45	0.37	n.a.	n.a.
21	Not politician	0.68	1.60	n.a.	n.a.
22	Peoples Interests	1.15	4.69	-1.11	-3.02
23	Policy for People	0.90	5.49	-1.14	-4.87
24	Predictability	-0.65	-0.54	n.a.	n.a.
25	Qualified	-0.99	-4.43	1.03	4.76
26	Racial	-0.62	-2.96	1.40	2.07
27	Religion	0.25	0.28	n.a.	n.a.
28	Respect	0.31	0.36	n.a.	n.a.
29	Sexism	0.31	0.81	n.a.	n.a.
30	Taxes	n.a.	n.a.	-1.91	-2.59
31	Trust	0.97	4.96	-0.95	-5.83
32	Women	-0.75	-3.47	0.19	0.20
33	Overall Preference	1.42	8.68	1.23	7.27

Table 5. Top performing items from table 1 (t-val >1.96)

Item	Words	*Trump*		*Clinton*	
		Coef.	t-val.	Coef.	t-val.
3	Capability	n.a.	n.a.	0.96	2.98
10	Experience	0.63	3.12	n.a.	n.a.
11	Guns	2.65	2.20	n.a.	n.a.
13	Integrity	n.a.	n.a.	-0.89	-4.23
15	Leadership	0.93	3.77	n.a.	n.a.
17	Likeability	1.02	5.36	1.19	5.40
18	Minorities	2.65	2.20	n.a.	n.a.
22	Peoples Interests	1.15	4.69	-1.11	-3.02
23	Policy for People	0.90	5.49	-1.14	-4.87
25	Qualified	-0.99	-4.43	1.03	4.76
26	Racial	-0.62	-2.96	1.40	2.07
30	Taxes	n.a.	n.a.	-1.91	-2.59
31	Trust	0.97	4.96	-0.95	-5.83
32	Women	-0.75	-3.47	n.a.	n.a.
33	Overall Preference	1.42	8.68	1.23	7.27

Trump has *positive* skews for (13) Integrity; (17) Likeability; (22) People's Interests; (23) Policies for the People; and (31) Trust. *Negative* skews for Trump are (25) Qualified; (26) Racial; and (32) Women.

Clinton skews *positively* for (3) Capability; (17 Likeability; (25) Qualified; (26) Racial; and (30) Taxes. *Negative* skews for Clinton are (13) Integrity; (22) Peoples Interests; (23) Policy for People; and (31) Trust.

In terms of Overall Preference i.e., Item 33 at the bottom of table 5, Trump scores higher than Clinton, i.e., Trump t-value = 8.68 and for Clinton the t-value = 7.27. Although these results are close, overall Trump is slightly ahead. This is our first real indication of Trump's edge over Clinton. However, from the word frequencies alone we are still lacking sufficient understanding of the key aspects driving this edge.

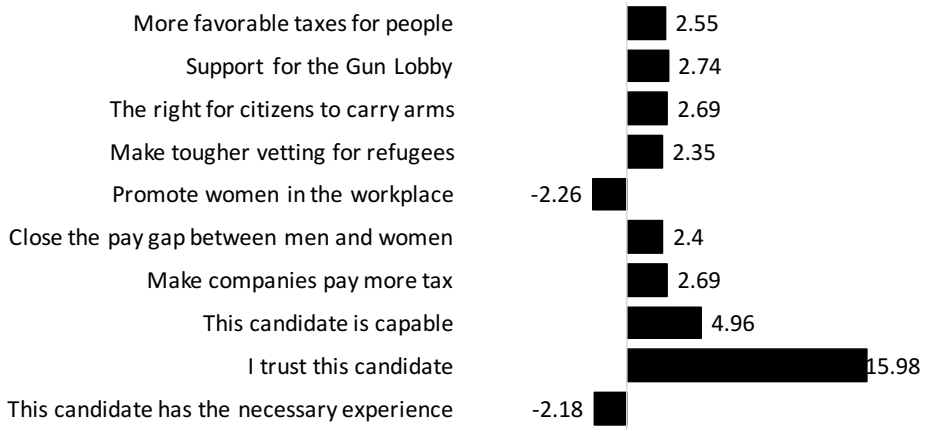
This is because when performing a content analysis on open text, trying to extract themes is a key challenge of text analytics programs. These problems are usually that of *valence* (positive or negative direction), *stop words* (filler words that are usually ignored such as 'the' or 'and') and *lemmatization* (the morphology of words e.g., 'cat' versus 'catty'), which either individually or in combination, make it hard for either the human coder or text software to make sense of what the subject was attempting to convey.

Quantitative enhanced data via suggested sentences

As we have discussed, the reason we provided respondents with suggested sentences immediately following their open text responses, was to attempt to retrieve factors or latent variables. We did this to capture the underlying dimensions that were driving their open free form text responses.

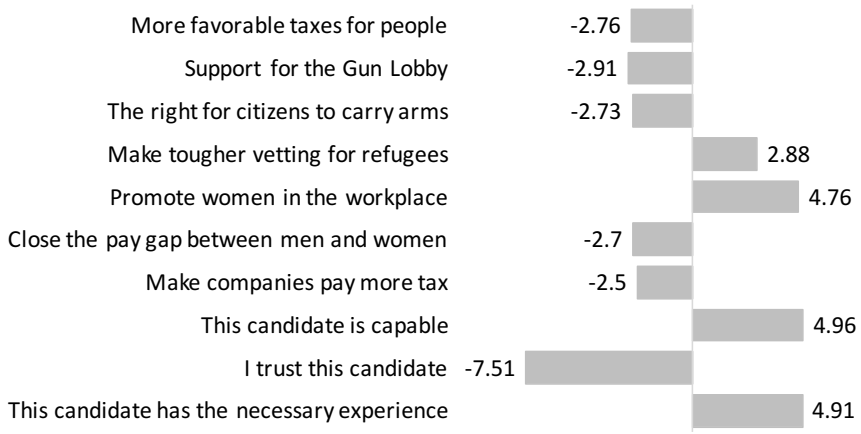
The first thing we noticed was that we had more distinct and intuitive information extraction from suggested sentences than from open free form text alone, where suggestions in real time made it easier to capture the true meaning of responses quantitatively, as shown in figures 5 and 6. In the case of single words, the valence of the word context can be called to question, depending on the software or coder that is used (Friedman, R. & Sanger, J., 2006; Kao & Poteet, 2010). Further, stop words make it harder to gain accuracy in machine or human coding (Konchady, 2008; Miner et al., 2012; Zanazi, 2007), whereas suggested sentences as in our case, leave little or no ambiguity, regarding the true valence the respondent is using.

Figure 5. Trump t-statistic scores from table 6



Looking at figure 5, Trump is viewed *positively* for items: 'More favorable taxes for the people'; 'Support for the Gun Lobby'; 'The right for citizens to carry arms'; 'Make tougher vetting for refugees'; 'Closing the pay gap between men and women'; 'Make companies pay more tax'; 'This candidate is capable'; and 'I trust this candidate'. Conversely, Trump performs *negatively* on 'Promote women in the workplace'; and 'This candidate has the necessary experience'.

Figure 6. Clinton t-statistic scores from table 6



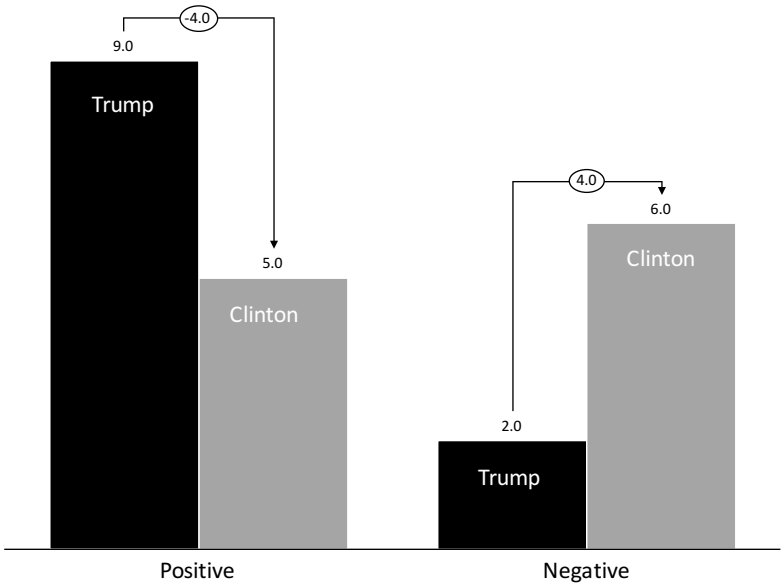
Looking at Figure 6, Clinton is viewed *positively* on; 'Make tougher vetting for refugees'; 'Promote women in the workplace'; 'This candidate is capable'; and 'This candidate has the necessary experience'. Conversely, Clinton performs *negatively* on: 'More favorable taxes for people'; 'Support for the Gun Lobby'; 'The right for citizens to carry arms'; 'Close the pay gap between men and women'; 'Make companies pay more tax'; and 'I trust this candidate'.

Note that with the enhanced quantitative study, the elements of *capability* and *trust* have emerged just like they did in the qualitative study (figure 3). Considering that capability and trust were the most mentioned points of differentiation by the media during the campaign (Roberts, 2016), the enhanced quantitative study appears to get us closer to the emotional drivers that turned the election (Collinson, 2016) just like the qualitative study did, by way of a 100 fold increase in sample size.

As an aside, in the qualitative study informants are guided as to the topic of conversation and imposed conversational structure by the moderator to help extract insights. The quantitative open free form text confirms whether the qualitative findings are indeed correct, because it is has no guidance imposed.

Note, using this approach there are eight positives and two negatives for Trump. Note also that there are four positives and six negatives for Clinton, as illustrated in figure 3. This is the second indication that Trump might have a lead over Clinton.

Figure 7. Suggested statements valence for Trump and Clinton



To further assess the outputs in figures 5 and 6, we performed factor analysis against the two candidates. As shown in table 6, we can see that Trump's overall message is perceived by respondents as largely comprising two latent variables or factors (K A Bollen, 1984; Kenneth A Bollen, 1989). LV1 groups elements of (1) trust; (2) capability; and (5) experience, as *personal* features of Trump. Also, LV1 groups elements of (3) toughness for refugees; and (4) the right to carry arms as emotive *campaign* features. These are the leading features of the Trump campaign. LV2 contains the more *rational* features.

Table 6. Higher order rotated components latent variable (LV) solution for Trump

	Trump		
Item	Item Description	LV 1	LV 2
1	I trust this candidate	0.97	n.a.
2	This candidate is capable	0.89	
3	Make tougher vetting for refugees	0.87	
4	The right for citizens to carry arms	0.86	
5	This candidate has the necessary experience	0.58	
6	Close the pay gap between men and women	n.a.	0.92
7	Make companies pay more tax		0.86
8	Support for the Gun Lobby		0.79
9	More favorable taxes for people		0.79
10	Promote women in the workplace		0.72

Using the same analysis for Clinton as shown in table 7, Clinton shares the same variables for LV1, but when we compare the loadings for Clinton to Trump, we see that LV1 for Trump is far more definitive. People know what Trump stands for.

Table 7. Higher order rotated components latent variable (LV) solution for Clinton

	Trump		
Item	Item Description	LV 1	LV 2
1	I trust this candidate	0.97	n.a.
2	This candidate is capable	0.89	
3	Make tougher vetting for refugees	0.87	
4	The right for citizens to carry arms	0.86	
5	This candidate has the necessary experience	0.58	
6	Close the pay gap between men and women	n.a.	0.92
7	Make companies pay more tax		0.86
8	Support for the Gun Lobby		0.79
9	More favorable taxes for people		0.79
10	Promote women in the workplace		0.72

Equally, for LV2, Trump leads Clinton in definitiveness of elements six through eight, meaning that Trump is solidly understood to represent these elements. Notice also that for Clinton, elements nine and 10 are further split into LV3 when compared to Trump, where they are both under LV2. This could indicate that Clinton is viewed as having two approaches to tax, one for companies and one for people. Trump on the other hand views tax holistically.

This approach supports the theory that being able to suggest sentences which the respondent can agree or disagree with, makes it more logical for the respondent to convey their response in a more meaningful way when undergoing open free form text exercises. How many times have we heard someone say when we paraphrase in conversation, the response: *"that is exactly what I was trying to say!"* We get the same effect by offering suggested sentences in open text data collection. It is as if we are paraphrasing the respondent based on their text response.

Suggested responses immediately following an open text question are used to confirm what the respondent was trying to say. A sentence such as *"More favorable taxes for people"* is more meaningful to both the respondent and the analyst. Similarly, *"This candidate has the necessary experience"* leaves little room for ambiguity for both respondent and analyst. It is easier to see the meaning of words when suggested sentences have been agreed upon by the respondent. An added benefit is that the provision of suggestions acts as a factor or latent variable.

When the suggested sentences acting as factors or latent variables have been factored via a higher order PCA, we can see that in nearly all cases, the correlations or factor loadings between the factors and the suggested sentences are much higher for Trump than for Clinton. This means that the clarity and connectedness and emotion that Trump supporters have is much higher than the supporters of Clinton. This is the third indication that Trump might have a lead over Clinton.

Tables 6 and 7 illustrate that in Trump's case, there are two higher order emotional factors that drive his supporters' understanding and interpretation of election campaign market information. In Clinton's case, there are three higher order factors. From this we can see that Trump's approach is easier to understand and takes less effort to interpret. We see this in the higher order factor loadings for Trump when compared to Clinton.

Conclusions

Summary findings

In this paper, we intimated that we would report findings from both a qualitative and quantitative study on subject matter that was high in emotional content. We used the qualitative and quantitative outputs from the qualitative study to contrast the *enhanced* form of questioning based on free form text from the quantitative study.

We proposed that the suggested sentences in the quantitative study offered an enhanced approach to gaining insights from open text, when offered to respondents immediately following their text input. We said that where there is a strong emotional content in what is being researched, that emotions would be captured using free form text just as well as they would be when using qualitative research, because open text is an unrestricted form of preference elicitation. We also discussed how open text offers the opportunity to be able to statistically determine the significance of themes with much larger sample sizes than can be realized with qualitative research, which usually leaves statistical significance indeterminable.

We talked about how higher order factors analysis for both candidates helped support the premise that Trump was dealing in a much simpler messaging strategy, based on emotion, rather than the rational foundation that Clinton was using.

In our case, the information that emerged was critical in the election outcome. The election at the end of the day hinged on two elements of *trust* and *capability*. Towards the end of the campaign, the media focused on the negatives of capability of an inexperienced Trump, and trust issues for Clinton. The enhanced open free form text solution by way of suggested sentences picked these two elements up just as the qualitative research did, in addition to other important elements for each candidate.

The attributes in the qualitative Likert-type scale questions were too generically emotive or practical, and indeed delved into all areas of the qualitative research outputs. This put pressure on informants to switch between their heuristic strategies i.e., qualitative reasoning versus quantitative reasoning in the case of emotional versus rational scales (Hegarty, Just, & Morrison, 1988; Tversky, Sattath, & Slovic, 1988). The open text suggested sentences did not suffer the same impositions.

We saw throughout our discussions that although the media portrayed Clinton as leading Trump throughout the election process, three clear instances emerged indicating that Trump might have had a lead over Clinton!

We have attempted to demonstrate that open text data collection can indeed benefit from suggested responses immediately following an open text response and do well in mimicking qualitative research and extracting emotional content.

We suggest that this is especially the case when the topic under investigation engenders emotion such as the recent US election. Respondents might struggle to express themselves in open text. The paraphrasing aspect of suggested sentences could do very well in cementing the true meaning of what the respondent wants to say in open text and offers insights in large sample sizes!

Implications for the Asian market

This breakthrough approach has far-reaching implications for the Asian market, with its high diversity in cultures, market sizes and market development. In particular, Asia's consumers originate from a multitude of ethnic and cultural backgrounds, and as a result their preferences and personal tastes are constantly evolving.

This has been previously illustrated in studies on magnitude estimation, where the verbal definition of *'excellent'* may not equate to the same extent on a 10-point rating scale, depending on the *'politeness factor'* embedded in the culture being surveyed. For example, verbal probing has revealed that an 8 out of 10 would be considered *'excellent'* in some cultures, whereas other cultures would require a rating of 9 or 10.

Similarly, when researching a topic where emotions are a key driver to consumer preferences, a standard quantitative survey approach with pre-coded attributes may be insufficient to adequately understand the key factors driving behaviour, both within and across a diverse cultural base. The proposed approach involving suggested responses immediately following open text responses allows the researcher to capture quantitatively those emotional factors.

To set up the proposed dynamic coding approach requires individual comprehension and interpretation of open text responses to develop the on-the-fly suggested responses, combined with machine learning linking likely response to the test being entered. As described earlier, the collaboration of human analysis and machine learning eliminates the need for an exhaustive text mining and analysis procedure, while also increasing the accuracy and speed of feedback to the clients. Moreover, the process provides us with an early and efficient indication of the full choice set of suggested responses, reducing cost and time.

This can be a critical advantage in the fast-paced Asian markets, where the frightening speed and scale of change in Asian consumer markets can dumfound and surprise even experienced researchers, presenting challenges in capturing the most relevant set of measures to include in quantitative research design. Being able to tap into emotions and a wider and more complete set of factors during the fieldwork period itself, could make for a significant difference in understanding key drivers of customer behaviour, and in capturing the rich, diverse and evolving consumer needs in Asia

Bibliography

Bollen, K. A. (1984). Multiple indicators: Internal consistency or no necessary relationship? *Quality and Quantity*, *18*, 377–385.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons Inc.

Collinson, S. (2016). *Emotions run high in final days of the campaign*.

Cosgrove, K. (2016). *The emotional brand wins*.

Darwin, C. (1875). *The expressions of emotions in man and animals*. Chicago: University of Chicago Press.

Eberl, J. M., Boomgaarden, H. G., & Wagner, M. (2005). One Bias Fits All? Three Types of Media Bias and Their Effects on Party Preferences. *Communciation Research*. http://doi.org/10.1177/0093650215614364

Ekman, P., & Friesen, W. V. (1975). *Unmasking the face*. Englewood Cliffs, New Jersey: Prentice-Hall.

Ekman, P. & Friesen, W. V. (1982). *Measuring facial movements with the facial action coding system*. Cambridge, England: Cambridge University Press.

Friedman, R. & SangerJ. (2006). *The Text Mining Handbook* (1st ed.). Cambridge University Press.

Friedman, S. M. (2012). Order and Quality Effects in Sequential Monadic Concept Testing: Methodological Details Matter in Concept-Testing Practice. *The Journal of Marketing Theory and Practice*, *20*(4).

Geer, J. G. (2006). In Defense of Negativity: attack ads in presidential campaigns. Chicago: University of Chicago Press.

Goldstein, W. M. (1990). Judgements of relative importance in decision making: Global vs. local interpretations of subjective weight. Organizational Behavior and Human Decision Processes, 47, 313–336.

Golshan, T. (2017). Donald Trump's unique speaking style, explained by linguists. Retrieved from http://www.vox.com/policy-and-politics/2017/1/11/14238274/trumps-speaking-style-press-conference-linguists-explain

Halvey, M., & Keane, M. T. (2007). An Assessment of Tag Presentation Techniques. In World Wide Web Conference. Alberta, Canada.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1985). Fundamentals of Item Response Theory. Newbury Park CA: Sage.

Hanrahan, C. (2016). Trump won US election thanks to the rural vote. Retrieved from http://www.abc.net.au/news/2016-11-10/donald-trump-us-election-rural-vote/8012874

Hegarty, M., Just, M. A., & Morrison, I. R. (1988). Mental models of mechanical systems: Individual differences in qualitative and quantitative reasoning. Cognitive Psychology, 20, 191–236.

High, S. C. (2003). Industrial Sunset: The Making of North America's Rust Belt. Toronto: University of Toronto Press.

Kao, A., & Poteet, S. (2010). Natural Language Processing and Text Mining. Springer-Verlag New York Inc.

Konchady, M. (2008). Text Mining Application Programming. Charles River Media.

Krippendorff, K. (2004). Content Analysis: An Introduction to Its Methodology. Thousand Oaks, California: Sage: Newbury Park, CA.

Likert, R. (1932). A Technique for the Measurement of Attitudes. Archives of Psychology, 140(55).

Malhotra, N. K. (2010). Marketing Research an Applied Orientation (6th ed.). New Jersey: Pearson Education, Inc., Upper Saddle River, NJ.

McGuffen, K. (2004). Political Attack Ads Can be Effective but Risky. Toronto.

Mercer, A., Deane, C., & McGreeney, K. (2016). Why 2016 polls missed their mark.

Mez Du, K. K. (2016). Donald Trump and Hillary Clinton, By Their Words. Retrieved from http://www.patheos.com/blogs/anxiousbench/2016/07/donald-trump-and-hillary-clinton-by-their-words/

Miner, G., Elder, J., Hill, J., Nisbet, R., Delen, D., & Fast, A. (2012). Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications. Elsevier Academic Press.

Myers, D. G. (2004). Theories of Emotion (7th ed.). New York, NY Worth Publishers.

NYT. (2016, July 28). Transcript: Hillary Clinton's Speech at the Democratic Convention. New York Times.

Odintext Inc. (2016).

Payne, G. (1986). Tom Bradley: The Impossible Dream (The Bradley Effect) (pp. 243–288). A Biography Roundtable Publication.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (2002). The Adaptive Decision Maker (2nd ed.). Cambridge: Cambridge University Press.

Reichheld, F. F. (2003). One number you need to grow. Harvard Business Review.

Roberts, D. (2016). Why Hillary Clinton lost the election: the economy, trust and a weak message. The Guardian. Retrieved from https://www.theguardian.com/us-news/2016/nov/09/hillary-clinton-election-president-loss

Sides, J. (2016). Is the media biased toward Clinton or Trump? Here is some actual hard data.

Skibba, R. (2016). Pollsters struggle to explain failures of US presidential forecasts. Nature, International Weekly Journal of Science.

SMH. (2017, January 21). US President Donald Trump's Inauguration speech transcript.

SMH. Sydney, Australia.

Stein, A., & Ettema, C. (2003). An overview of spatial sampling procedures and experimental design of spatial studies for ecosystem comparisons. Agriculture, Ecosystems and Environment, 94(1), 31–47.

Thurstone, L. L. (1927). A law of comparative judgement. Psychological Review, (34), 273–286.

Thurstone, L. L. (1928). Attitudes can be measured. American Journal of Sociology, (33), 529–554.

Tinseley, H. A., B. S. E. (2000). Handbook of Applied Multivariate Statistics and Multivariate Modelling (1st ed.). Academic Press.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgement and choice. Psychological Review, 95, 371–384.

Vijayarani, S., Ilamathi, J., & Nithya, K. (2015). Preprocessing Techniques for Text Mining - An Overview. International Journal of Computer Science & Communication Networks, 5(1), 7–16.

Zanazi, A. (2007). Text Mining and its Applications to Intelligence, CRM and Knowledge Management. (WIT Press, Ed.).

The Authors

Brian Fine is Chairman, Quality Online Research (QOR), Australia.

Con Menictas, Newcastle University, Australia.

Harnessing the Power of Voice

Analysing audio responses to detect excitement, passion and activation

Saurin Shah • Roland Leung • Geeta Lobo

Introduction

Marketers have always known the role that emotions play in consumers' product purchase process. Today, in addition to injecting emotions into new products through advertising, marketers are building emotions into their product designs. Researchers have recognized the importance of measuring emotions embedded in products but the traditional metrics of rating scales fail to reflect the true picture. Various neuropsychological technologies have emerged but none is easily scalable or easy to interpret. Voice analytic, which comprises linguistic and acoustic components, provides a highly viable new solution to marketers' needs. This paper illustrates the potential for mass adoption of voice analytic with a pilot study recently completed in India, illustrating how the analysis of consumers' spoken responses bring a level of insight that traditional rating scales cannot do alone.

Functional needs vs. emotional needs

As marketers and researchers, we know that products and services do not simply fulfil consumers' functional needs but also their emotional needs. For example, Rolex watches '…*do not just tell time, they tell history (witness moments of significant achievements)*'. And Google's search engine does not simply give you what you asked for, it expands the horizon of your knowledge and makes you feel enriched.

Some argue that consumers purchase products for both what the product does for them (i.e. product functions) and how it makes them feel (i.e. product emotions). Some researchers further maintain that products and services may range from high to low on the emotion scale but there is a consensus that emotion does form part of the overall experience. For example, the choice of gas supply service at your home may be high on functional aspects and very low on emotions. Yet a pop concert featuring your favourite artist will satisfy little functional need but be extremely high on emotional satisfaction (Boatwright & Cagan, 2010)

Role of emotions in consumer decision making

When products are laden with emotions that endure its lifetime, customers benefit as the product that exudes an on-going stream of positive emotions promotes a 'feel-good' aspect, reinforcing customer satisfaction with the purchase, tangibly improving their lives and increasing the likelihood of receiving the best form of advertising: word-of-mouth.

Emotions play a critical role in consumer choice and purchase decisions as well. In fact the role of emotions in intuitive decision making has been elaborated by Kahneman & Tversky (Kahneman & Tversky, 1985). Kahneman has offered a framework for understanding the two types of decision-making describing it as System I and System II thinking. Specifically, System I is quick, intuitive, emotional while System II is slow, deliberate and rational. System II typically comes into play if the decisions are unfamiliar or out of the ordinary. Our past experience and emotions bias our decisions towards the use of System I (Kahneman, 2011).

Much of consumer decision making is also likely to have a strong component of System I thinking. Conventional means of investigation into consumer behaviour rely on enquiry which is more likely to access System II thinking as the assumptions underlying most of the enquiry is that these decisions are largely cognitive, deliberate decisions.

How can intuitive responses be measured in the context of consumer decisions?

Knowing that emotions are important for products and services, our task as researchers is then to measure it. But that presents another challenge because the term emotion has meant different things to different individuals, be they scientists or laymen alike. Scientists who made serious attempts in defining it have broken it down to minimum of five components, including: cognitive (appraisal), neurophysiological (bodily symptoms), motivational component, motor expression (facial & vocal expression) and subjective feeling (emotional experience). When it comes to the measurement of emotion in an experimental context, scientists have resorted to the measurement of valence (positive-negative), arousal (calm-excited), and other dimensions that are relevant for the context (Scherer, 2005).

Implicit measure of emotions

The majority of traditional, mostly explicit methods of market research (e.g. survey using rating scales) address System II way of thinking. Explicit measurement methods assume that consumers have the ability and motivation to divulge the required information (Cunningham et al., 2001). When surveying emotions this is highly questionable, as the explicit methods are not in a position to access the intuitive and emotional System I way of thinking. This prompted discussions about alternative methods of measurement (Bagozzi et al., 1999). Implicit measurement in particular presents a sensible alternative to detect unconscious intuition and emotion. These methods can overcome the limitations of explicit methods in relation to the ability and motivation of consumers (Nosek et al., 2011).

Many of these implicit measurement methods are physiological and neurophysiological processes. These include the analysis of heart rate, brain activity, facial expressions, or language. Most of these methods can measure emotions, but mainly non-directionally, i.e. without explaining the trigger and direction of the emotion. Moreover, many of these methods are very complex and insufficiently scalable within the context of quantitative data collection.

Speech analysis therefore assumes a special position. The spoken word of consumers informs us simultaneously about conscious, rational aspects (what is said) and subconscious, emotional aspects (how it is said), thereby making emotion interpretable and enabling recommendations for action. Our approach utilizes this form of analysis within the scope of efficient and cost-effective quantitative research.

Model development

A classification model has been developed for the speech analysis approach, which: (i) examines transcribed verbal inputs in terms of their sentiment and activation level, and (ii) determines the speaker's level of emotion (passion) on the basis of different facets of their voice (Eyben et. al, 2017).

All parts of the classification model which are used have been developed according to the latest scientific standards. For this purpose, a variety of voice recordings (in German, English, Spanish, and Mandarin) and their transcripts were collected and manually coded. The result is the reference for subsequent analysis. These manually coded elements were used to train the model. In this context, it is important to note that all of the voice recordings were obtained during new products tests with consumers, i.e. the development of the classification model itself was based on the intended application of the approach and takes the specifics of the consumer language for this specific subject into account.

Only some of the manually coded elements were used to train the classification software. The software has been trained to reproduce the results of the manual coding. Some of the ground truth was used to test the accuracy of the classifiers on the basis of a "natural" sample that was not used for the training. The classification algorithm was then interactively optimized to increase the hit rate.

During the survey in which speech analysis was piloted, respondents were encouraged to provide spontaneous spoken feedback about the product or service concept that’s being presented to them. Voice recording utilizes the built-in microphone of the device during the survey (laptop, tablet, smartphone, desktop) - for the respondent, the experience is therefore no different from an ordinary online survey apart from the voice recording. The raw data collected during the survey is then available in the form of

numerous voice recordings, each one contains the non-guided, spontaneous feedback from a respondent about one of the test concepts.

The voice recordings are then evaluated according to two analysis paths: (i) *Linguistic Analysis* which focuses on what consumers say and (ii) *Acoustic Analysis* which focuses on how they say it.

Linguistic analysis

The voice recordings are transcribed to facilitate the analysis of the content of the respondents' feedback. This is currently performed manually, as automated speech-to-text methods can only deliver the text accuracy required in a few languages, if at all.

In addition to the wide range of unstructured respondent feedback which is available and qualitatively analyzed in text form after transcription or can be manually coded in the same way as conventional open responses, proprietary classification software provides two output variables: (i) *Sentiment***:** Is the content of the feedback about the concept positive, neutral, or negative? And (ii) *Activation***:** Does the respondent's feedback contain statements that indicate actions such as trying out the product or wanting to learn more about it?

Unlike other classification tools which are used, for example, in social media analysis, the models used in our approach are specially trained for concept feedback. This means that the data basis (ground truth) used for the modeling and optimization is based exclusively on manually coded consumer feedback about innovative product and service concepts.

Acoustic analysis

The second analysis path, audio signal analysis, uses classification software that was developed in cooperation with the 'Complex Intelligent Systems' chair at a German university. The software uses more than 1,000 characteristics/features of the human voice to analyze the speaker's voice and also tells us the excitement level (*Passion*) of each respondent and for each concept as a further data point.

The software also ensures that the efficiency of the survey method is as broad as possible, i.e. the usability of the audio signal is largely unaffected by ambient noise. A passing bus, for example, can be reliably filtered out.

Relevance of the new approach to consumer investigations in emerging markets

An approach that offers access to the intuitive feelings of consumers and yet is simple to administer has special advantages in the context of emerging markets such as India. Conducting innovation research in such markets, especially for categories with mass appeal, poses challenges at multiple levels.

Specific challenges of conducting innovation research in large emerging markets

In markets such as India field investigation for quantitative studies deploys the services of a large body of freelance investigators, spread over the major cities across the country. The challenges of representing a vast heterogeneous population calls for large sample studies to be conducted in multiple locations, concurrently. In India in particular, the field investigation is now unable to attract the skilled personnel that it requires mainly due to emergence of multiple more remunerative job alternatives in the service sector. Conducting fieldwork with modestly skilled investigators has accentuated the following problems and has made the need for an alternative investigation methodology even more pronounced.

Conventional innovation research such as product idea screening, concept testing, product testing, concept usage test and volumetric estimations, rely on aided enquiry captured on standard response scales. A set of parameters led by purchase interest, along with relevance, uniqueness and credibility, form the mainstay of such investigations. The problems posed by this conventional approach are summarized below:

1. *High correlation among the parameters*: Though in principle most of the parameters included represent independent dimensions of product evaluation by consumers; there tends to be very high correlation between the various parameters being captured. In most research concepts either uniformly score well on most parameters or fail the norms across all. This lack of discrimination implies that performance evaluation of concepts become rather unidimensional.

2. *Low reliability of response scales*: The scales used in these investigations especially translated into multiple regional languages, have proven to be low on reliability. Most of these are not even true interval scales and 'research on research' studies in the past have repeatedly shown that it is difficult to get consistent results from repeated experiments because of these scales.

3. *High top-boxing and consequent low variance in responses*: Cultural factors lead to high top boxing in many of these markets, with consumers reluctant to openly express adverse or even neutral opinions. This effectively reduces most of these 5 point scales to a 3 point scale bringing down ability of the scale to discriminate between concepts and product ideas even further.

4. *Poor quality of unaided responses collected*: Given the scale and spread of these quantitative studies conducted in multiple languages, even though there are standard open ended responses included in the questionnaire, the modest quality of investigation implies that these are limited, poorly recorded and not accurately translated for coding and tabulation. The transmission losses in the process alone bring down the value of these responses to the research. For most quantitative innovation research, getting access into intuitive or even natural responses of consumers is quite difficult.

5. *Long and fatiguing investigations*: The inefficiency of scales and the drawbacks of the other enquiry is sought to be compensated through added diagnostic questions. These include long attribute batteries and directed open ends all of which add to the length of investigations making the enquiry even less reliable due to high respondent fatigue.

6. *High non-sampling errors*: Dependence on a vast body of briefly trained investigators adds to the problem because it is impossible to neutralize the effect of the individual style of investigation and engagement on the responses collected. The non sampling error due to variation in investigation style is believed to be substantial especially for large multi-centre studies.

Focus of the voice analytic pilots in India

The purpose of our investigation was to examine if using a methodology that combines linguistic and acoustic analytics can effectively address these problems. Specifically we had the following expectations from the investigations.

- Does the inclusion of intuitive responses to rationalized reactions provide better differentiation between product ideas evaluated?
- Does this approach offer a more discriminating analysis than the conventional scale guided approaches?
- Does this approach enhance the diagnostic value of innovation research?
- Can this simple enquiry effectively replace scale based investigations?

In a world of ever increasing clutter and shorter attention spans, working out which of the product ideas has the potential to succeed is rapidly becoming the main challenge for most marketers. We had a set of clients who combat with this problem on a daily basis and they joined us in this quest for a better innovation evaluation.

Specifics of the pilot study

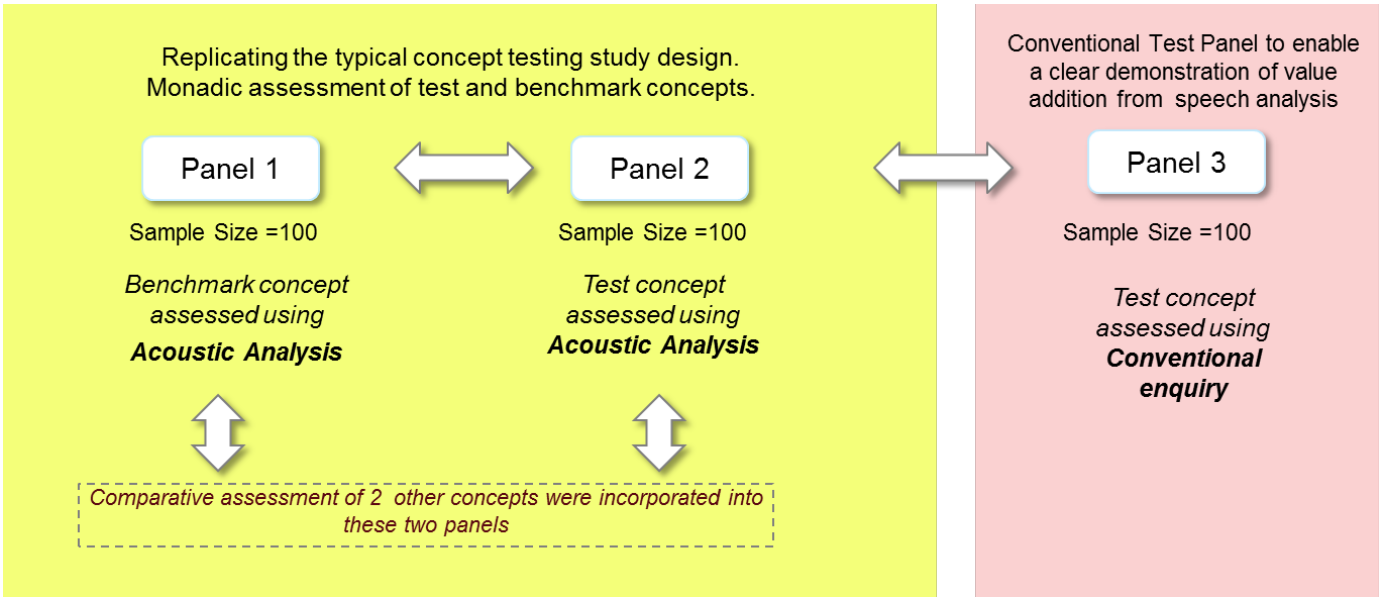
We included new concepts from food and personal care categories into this investigation. In all six concepts were tested. The leading FMCG players in India – Nestle India, Britannia Industries, L'Óreal and Godrej Consumer Products – collaborated in this pilot effort.

Since the acoustic coding for English has been validated and proven to be culturally neutral, for this pilot study we focussed on offerings that would be relevant for the higher affluence metro consumers, who could converse comfortably in English.

Three matched consumer panels were created. In two of these panels we used the newly developed acoustic analysis approach and in the third we used a conventional concept investigation.

The two panels for acoustic analysis mirrored the typical concept investigation design followed, viz; monadic evaluation of test and control concepts. The third panel had the test concept tested using conventional enquiry. Figure 1 illustrates the study design in detail.

Figure 1. Study design in detail



Each panel included 100 consumers from Bangalore and Delhi (n=50 from each city). In all, n=300 consumers participated in the pilot study. As given in the panel design details, the acoustic analysis panels evaluated two more concepts in addition to the monadic concept.

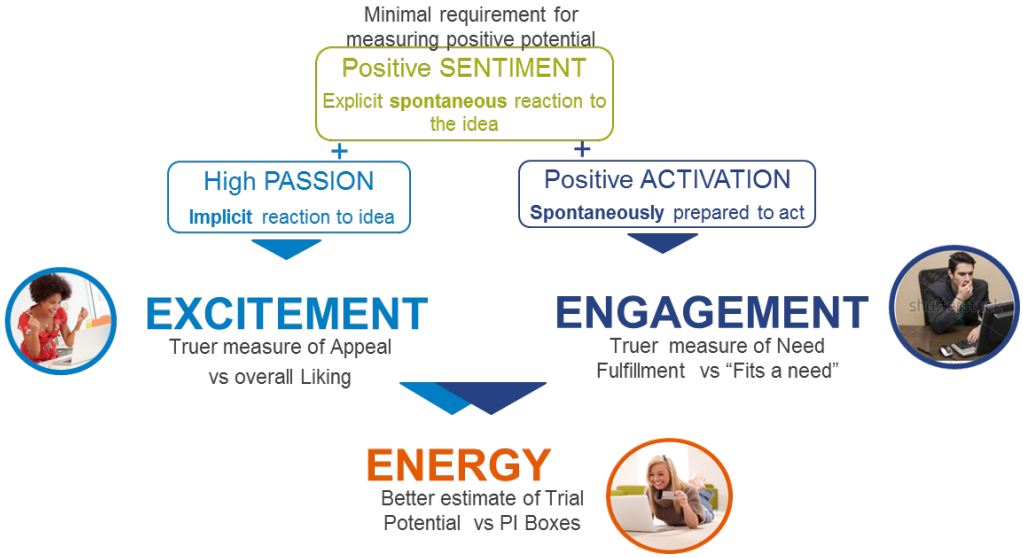
Concept investigation

The enquiry for the acoustic analysis panels consisted of a simple question *"What do you think of this product description and why?* There were no scales or any other aid to response. Consumer responses were completely spontaneous and unprompted. They were merely encouraged to speak for 10 seconds in response to this question, after having read the product description.

The response metrics

We derived three response metrics from this simple spontaneous enquiry, namely: (i) Sentiment, (ii) Activation, (iii) Passion.

Figure 2. Key metrics derived for the acoustic analysis panels



These three measures taken by themselves and in conjunction were used to assess the response to the product ideas or concepts. The core metric was one which looked for signs of positive scores on Sentiment, Activation and Passion, if the response was positive on all these three measures, we coded it as "Energized".

Key findings

The responses were analysed not just to assess concept performance but to see if this novel approach could help us arrive at a more robust measure of consumer interest in product ideas. The results are very encouraging and there are four aspects on which this approach seems to offer an improvement over the conventional scale based enquiry.

1. Does inclusion of intuitive responses provide better differentiation between product ideas evaluated?

Figure 3. Key metrics derived for the acoustic analysis panels

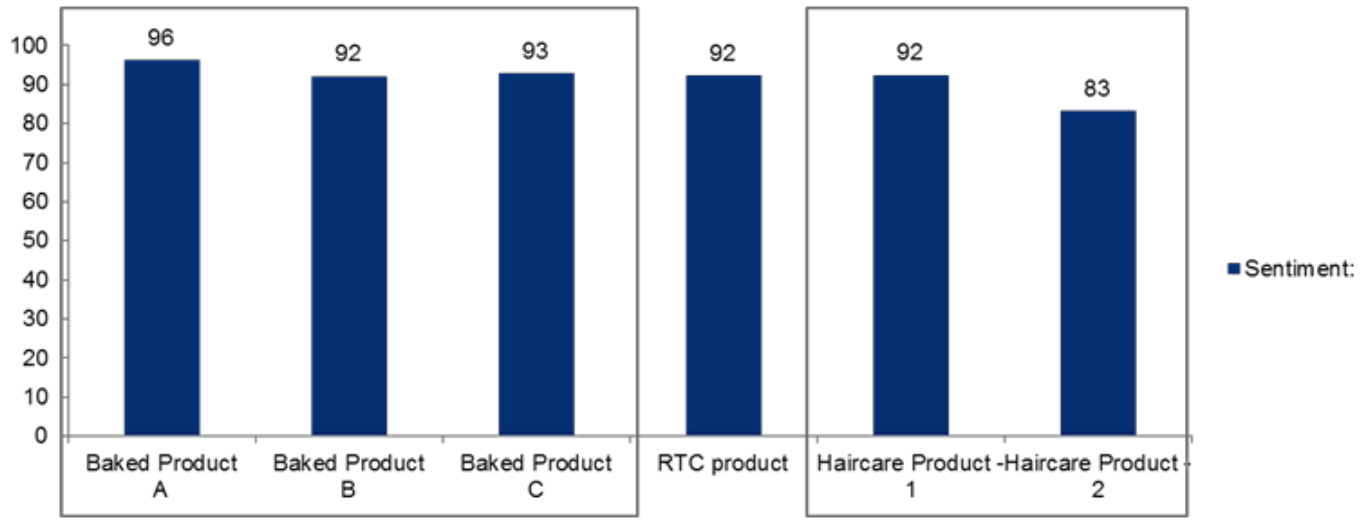
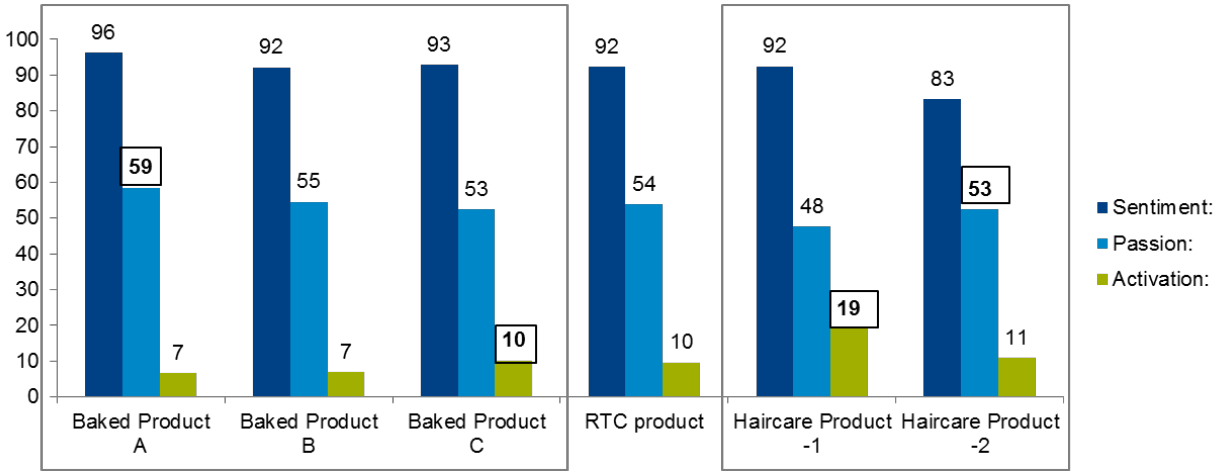


Figure 3 indicates that when assessed in terms of Sentiment alone all the concepts score high. This is also the reason why conventional measures which rely on stated disposition do not offer the best differentiation between product ideas.

Figure 6. Passion and Activation scores



The Passion coding as is evident from figure 4 has emerged as an independent measure. The ideas with the highest Passion score are not always the ones where the consumers stated disposition – reflected in Sentiment score – is the highest. or their inclination to act – reflected in the Activation score – is the highest. Acoustic analytics are clearly contributing a new dimension to understanding consumer responses.

2. Does this approach offer a more discriminating analysis than the conventional scale guided approaches?

Of the concepts included in the pilot study, we had tested one concept through conventional monadic scale based assessment as well. For this concept the results in table 1 indicate that the conventional approach was characterized by straight-lining of responses and by high top-box scores across measures. The three parameters of the present approach are more discriminating.

Table 1. Comparison of results from conventional enquiry and the voice analytics based enquiry scores

Baked Product -C			
Conventional Enquiry *(Top 2 boxes %)*		**New Approach %**	
Purchase intention	93	Net Sentiment Score	93
Relevance	80	Engaged	10
Uniqueness	69	Excited	50
Credibility or Believability	78		

3. Does voice metrics help access intuitive or implicit consumer responses?

At the heart of the approach is the belief that through the acoustic measures we will be able to tap into the intuitive consumer responses, described as System I thinking by Kahneman (Kahneman, 2011). This is also information that cannot be derived or gauged from the stated or scale based responses, which tend to reflect the cognitive or System II thinking, nor can these be obtained from the open-ended responses which tend to be prone to multiple response quality problems.

In order to address this question, we examined the difference in the stated responses among those who had a positive Passion score as compared to those who only had positive Sentiment and Activation score. In the present system the measure *'Engaged'* refers to those consumers with a positive Sentiment score

(an expressed or explicit favourable spontaneous reaction to the idea) as well as positive Activation (those who have spontaneously indicated an inclination to act). '*Energized'* refers to the *'Engaged'* consumers whose implicit responses as measured by voice metrics is also positive.

Table 2. Assessing the incremental contribution of Passion coding

% of response mentions	**Engaged but not Energized**	**Energized**
Hair-care product idea		
Sensorial positives	8%	**16%**
Benefits/ Ingredients	16%	20%
Baked Foods product idea		
Sensorial positives	3%	**9%**
Benefits/ Ingredients	35%	20%
Specific consumer mentions	2%	**8%**

Comparison of the stated responses among these two groups indicates that the Energized did not necessarily mention a higher number of benefits for the product ideas. The key difference was that they spoke more about the sensorial elements of product consumption or usage. There was also the evidence that for some concepts, the Energized spoke more about specific family members consuming the product.

Sensorial impressions are not cognitive. The ability to visualize consumption or usage moment is also more reflective of intuitive response to an idea. Thus the incremental contribution of the Passion score, as reflected among those Energized, seems to be identifying those consumers who have intuitively warmed up to a product idea. These findings do indicate that the Passion score from acoustic analysis seems to reflect intuitive responses not accessible through conventional investigation.

4. Does the new approach have operational advantages?

The enquiry is extremely simple and does not tax the respondents at all, who merely have to respond to a direct question. This single response is used to derive all the metrics.

The time taken to administer a single concept response is 5-7 minutes, as compared to a 15-25 minutes enquiry required for administering multiple scale based questions.

While the length of enquiry is substantially shorter, both the derived measures and the stated responses using this approach yielded rich diagnostic material. In fact through the 10 second response to the basic question we were able to gain more content than what is obtained through conventional open ended enquiry, which typically yields two to four coded responses per consumer.

Applications of this methodology

The fundamental strength of this approach is in that it provides easy access to the implicit consumer responses that drive consumer choices. The acoustic analytics methodology was developed for concept testing, but these early experiments strongly suggest application of this approach to a wider range of NPD research and innovation work.

As a method that allows us to access the implicit consumer responses and combine these with their cognitive responses, these simple metrics (Passion, Sentiment & Activation) and their derivatives (Engagement, Excitement, Energy) can readily be used in any research where conventional scale based enquiry is being deployed currently. These include:

- Benefits screening
- Early idea testing
- Concept testing
- Product testing
- Concept usage tests
- Full offer tests

This approach can also replace the more cumbersome and intrusive approaches to accessing implicit responses, that rely on physiological and neurophysiological measurements and require specialized interpretative skills. Among industries, implicit consumer responses are likely to have a stronger play in predicting consumer behaviour in contexts where sensorial impressions are crucial. This includes almost the entire consumer goods sector as well sectors like Technology and Media where rationalized responses are likely to be less influential in predicting behaviour than implicit responses.

Conclusions

The present study holds out promise for an improved enquiry for innovation research especially concept and early ideas screening. The findings are very exciting and we plan to extend this approach to different stages of new product development process. Further work is needed to establish the validity of this method. But the advantages evident already make a strong case for pushing this approach as a credible alternative to conventional scale based consumer testing.

References

Peter Boatwright & Jonathan Cagan, 2010, Built to Love: Creating Products That Captivate Customers, Berrett-Koehler Publishers, Inc.

Amos Tversky & Daniel Kahneman, 1985, The framing of decisions and the psychology of choice, Environmental Impact Assessment, Technology Assessment and Risk Analysis, Springer Berlin Heidelberg

Daniel Kahneman, 2011, Thinking Fast & Slow, Farrar, Straus and Giroux

Klaus R Scherer, 2005, What are emotions? And how can they be measured, Social Science Information, SAGE Publications

Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. Journal of the academy of marketing science, 27(2), 184-206.

Cunningham, W. A., Preacher, K. J., & Banaji, M. R. (2001). Implicit attitude measures: Consistency, stability, and convergent validity. Psychological science, 12(2), 163-170.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. Trends in cognitive sciences, 15(4), 152-159

F. Eyben, M. Unfried, G. Hagerer, and B. Schuller, “Automatic Multi-lingual Arousal Detection from Voice Applied to Real Product Testing Applications,” in Proceedings of ICASSP, New Orleans, LA, March 2017. 5155-5159

The Authors

Saurin Shah is Head of CMI, Godrej Consumer Products Limited, India

Roland Leung is Commercial Director, Market Opportunities & Innovation, APAC, GfK, Hong Kong.

Geeta Lobo is Senior Vice President, Consumer Experiences, GfK, India..

Uncovering the Unspoken Word

How implicit mobile research techniques can minimize response biases in Asian markets

Paul Janssen • Robin De Rooij • Marina Georgieva

Introduction

To drive growth, Unilever and other multinational companies have their eyes set on quickly growing Asian markets. Gaining a deep understanding of consumers is a pivotal part of realizing that growth. However, when doing research in this region it can be difficult to reach the right people and get a true understanding of their preferences. Many Asian cultures are characterized by high context effects, which means that in social settings many things are often left unsaid. Asian cultures are also characterized by higher degrees of acquiescence and embeddedness, which can lead to Acquiescent Response Style (ARS) and Socially Desirable Responses (SDR). ARS refers to the tendency to agree with propositions in general regardless of the content. While SDR is defined as the tendency of respondents to answer questions in a manner that will be viewed favorably by others, ARS and SDR make it more difficult to uncover the real opinions of respondents. When explicit judgements about stimuli are requested, these effects are particularly strong and they are further strengthened when an in-person interviewer is present. We believe that in order to overcome these response biases, we need to adjust and move towards a different way of conducting research, one that is more in line with current times.

Dual processing

While accurately measuring consumer behavior is particularly difficult in Asian markets, it is something all markets struggle with. In an ever-changing and digitalized world, marketing and research professionals are constantly challenged to find new ways of interacting with consumers in an effort keep up with changing attitudes and preferences. During the past two decades, an important compilation of neuropsychological research has deepened our understanding of how the human brain actually works and raised some serious questions about universally accepted research techniques. Traditional research methodologies predominantly activate cognitive processes in the brain, creating a bias toward more rational outcomes. Think of the artificial and cognitive nature of rating, ranking and MaxDiff exercises. These exercises tend to make respondents overanalyze and gravitate towards rational outcomes that are not fully reflective of actual consumer behavior that is also heavily influenced by subconscious processes in the brain.

Around 2011 when Daniel Kahneman published his famous ‘Thinking Fast and Slow’, the market quickly gained a better understanding about the infamous systems 1 and 2 and things slowly started to change. This so called dual processing of information theory is also an important theoretical ground on which the new research techniques that will be introduced later are based. For decades psychologists, economists, and behavioral scientists have researched the brain’s two distinct modes of processing information, labeled System 1 and System 2. System 1 refers to the brain’s fast, automatic and intuitive side and System 2 refers to the mind’s slower, analytical mode where reason dominates. The two systems do not operate in isolation and are often in interplay with each other. For example, the sound of a gunshot activates System 1 involuntarily, but the brain’s attention soon shifts to System 2, where the information will be assessed at a more conscious and analytical level. Although System 1 is considered dominant, considering either system in isolation risks overlooking important elements of the decision-making process as both systems work in close collaboration with each other.

Bridging rational and emotional drivers

While traditional research methods are good at measuring rational or conscious processes of the brain, implicit research techniques are effective in measuring the subconscious processes. The Implicit Association Test and Effective Priming are the two most commonly used implicit techniques and were designed to detect the strength of a person's automatic association between concepts (e.g., black people, gay people) and evaluations (e.g., good, bad) or stereotypes (e.g., athletic, clumsy). In a commercial setting, these tests can be useful in measuring attitudes and beliefs towards products and brands that people may be unwilling or unable to express. However, some controversy still exists about what these techniques actually measure, which limits the actionability of the results. Their commercial use is further limited by the large sample sizes they require and the tedious and lengthy nature of the exercise. With these considerations in mind SKIM's objective was to create a new methodology that not only bridges conscious and subconscious drivers, but is also practical and actionable in nature.

The new mobile reality

In addition to a heightened awareness about the neuro-psychological processes outlined above there are other key developments that should inform the development of any new research method in this day and age. Mobile devices are now at the center of our daily lives, connecting us 24/7, and their dominance is only set to rise. By 2020, an estimated 70% of the global population will own one and 50% of online sales will be made via a mobile device. The use of mobile technology is particularly relevant for the Asian market. In most countries, mobile internet access has leapfrogged wired access. For instance, in Thailand over 70% of the people access the internet more frequently through mobile devices than desktops. In China, nearly 90% of people with internet access will connect to the internet through their phones.

For research professionals this provides problems and opportunities at the same time. It means that many existing methods lose their relevancy quickly. On the other hand, embracing mobile as the main form of communication provides an opportunity to interact with consumers in a way they are most comfortable with. In many cases it will also eliminate the need for an interviewer and can help us reach people that are underrepresented in traditional online panels. Although we recognize that mobile panels are still in a developmental stage, all statistics point towards mobile as the best solution to connect with truly representative samples of consumers in the future. The rise of mobile technology is also a great contributor to the world of instant gratification we live in today, where needs are immediate and even a download delay of two seconds is enough to abandon a website, a world where we can search in a minute, order with a click, like with a swipe and share in a moment. To stay ahead and deliver on customer expectations in these changing dynamics we, as marketing professionals, must re-think established processes, change the way we measure, and find new, more relevant and engaging ways of interacting with consumers.

New research engine

With these key trends in mind, our objective was to develop an engaging mobile first research engine that captures conscious and subconscious processes of the human mind and can be used across different application areas. To achieve this goal a cross-functional team of psychologists, brand communication specialists and the world's leading discrete choice modeling experts was formed. The methodology that was eventually created bridges the gap between traditional and implicit techniques by using intuitive and engaging exercises that trigger more instinctive responses and incorporating reaction time as an implicit measure. Advanced algorithms rooted in SKIM's choice modeling expertise use the recorded choices and reaction times to model outcomes that are more representative of the true behaviors consumers display in market. More specifically, the methodology consists of two core modules.

Attraction module

In this module consumers repeatedly swipe stimuli to the right if they like it and to the left if they do not. The exercise is based on quick intuitive judgments made one right after another. These judgments are made in isolation, as we know that these judgements tend to be more emotional, automatic and associative. This approach is particularly good for measuring initial reactions to stimuli, for example initial appeal or identification of which stimuli break through the clutter. The approach is well suited to test a variety of stimuli, including print ads, messages, product ideas, digital banners and pack shots. Stimuli are scored on a scale from 0 to 100, calculated on the basis of the swipe direction and normalized reaction times. The decision to use likeability as the dimension along which respondents have to swipe was a carefully considered one. System 1, which is also known as the "heuristic" system, has a strong connection with the experience of affect. In psychology affect is the experience of feeling and emotion. Affect can be expressed in many ways, but at its core it is simply based on how much we like or dislike something. Towards things we like we develop positive feelings and emotions, and towards things we do not like negative ones. Because of its close ties to the dominant system 1, affect is an important behavioral heuristic. This in turn makes likeability, especially when asked in an intuitive and engaging way, a good predictor of actual consumer behavior.

Conversion module

In the conversion module, consumers make trade-offs between two competing stimuli. Similar to the conversion module a wide variety of stimuli can be tested. However, we have to keep in mind that real estate on a mobile screen is limited, especially when presenting two stimuli at the same time. By making consumers choose between options, joint judgments about the stimuli are created, which tend to be more rational in nature than judgments made in isolation. This exercise is useful when simulating moment-of-truth situations where choices have to be made and more rational and deliberate processes tend to kick in, like at the supermarket shelf or in an online shopping portal. Responses are analyzed using a discrete choice model that uses the choices and reaction times as inputs. By using an adapted version of Hierarchical Bayes (HB) that also takes reaction times into account, consumers' sensitivities (utilities) towards the different stimuli are estimated. These utilities then feed into a simulation model that calculates the brand's share of preference increase or decrease versus competition generated by each of the tested stimuli. Discrete choice models are not new and because of their ability to generate derived rather than self-stated results have always been a popular method to reduce response biases. The mobile conversion module outlined here minimizes response biases even further. First, by having only two choices on the screen and using a simple screen tap as the indicator of choice, the exercise becomes faster and more engaging than a traditional discrete choice model. This helps activate System 1 in addition to System 2, capturing both conscious and subconscious processes. Including reaction time as an implicit measure in the utility estimation process enriches the model even further.

Cross-country differences

Research setup

To determine how the new mobile research methods affect differences in response biases that exists across Asian cultures, an ad hoc cross-country research with a sample size of N=600 per country was conducted. The research covered four countries. India, Philippines and Singapore represented a set of Asian cultures where response biases were expected to be present and Australia served as representation of a more western culture where response biases are expected to be less significant.

The research methods tested were a standard rating scale, Maximum Difference Scaling (MaxDiff) and the attraction or 'swiping' module of the newly developed mobile research technique. The rating and MaxDiff cells were conducted through a typical online survey environment, while the 'swiping' cell was conducted on smartphones. (Both cells were recruited through the same panel and assigned based on the device

used to access the survey). Product claims for the shampoo category were used as test stimuli to compare the three methods. In each of the countries the following 10 claims were tested.

Table 1.

1	Salon style results at a fraction of the cost
2	The original shampoo from <research country>, for smooth and beautiful hair
3	No.1 bestselling shampoo in <popular foreign country>
4	48 hour freshness. Forget about smelly and oily hair!
5	Give your family the beautiful hair they deserve
6	Feel like a superstar with our most luxurious care yet
7	Stand out! Hair so beautiful, even your friends will be jealous
8	Forget the rest, research shows our shampoo performs best!
9	Softer, smoother hair after just 1 wash
10	95% less split ends for smooth and beautiful hair

The shampoo category was used because of its universal appeal and high penetration across markets . The claims were selected to ensure a good representation of different product characteristics that could trigger different responses across cultures. Lastly, claims were tested in English, which is a commonly spoken language in all countries. This was done in order to keep the results as comparable as possible between countries and rule out any noise that could result from translations of claims.

In order to compare the level of Socially Desirable Responses across the three methods, a proxy for social desirability of each of the claims was included. This was done by measuring how important each of the claims' underlying product characteristics was in each of the countries. An overview of the matching product characteristics per claim can be found below.

Table 2.

1	Is affordable
2	It is an <own country> brand
3	Imported from <popular foreign country>
4	Eliminates bad smells
5	Suitable for the entire family
6	Is a luxurious product
7	Helps you stand out
8	Is better than other shampoos
9	Gives results quickly
10	Reduces split ends

Result analysis

Three types of analysis were carried out, each with a different purpose. First, average scores across methods were compared to established if the new more intuitive mobile approach could provide a better base of comparison across countries; with comparison being especially difficult when Acquiescent Response Style bias is more prominent in some markets than others. Second, the relationship between the social desirability of the claims and their performance across the different methods was analyzed to determine if the new method could reduce Socially Desirable Responses. Third, respondent feedback helped to established engagement levels for each of the methods.

Cross-country comparisons

The average scores per method, including an index number in parentheses can be found in the table below. Please note that MaxDiff scores are based of zero-summed utilities, making them relative in nature and not suitable for an analysis of absolute scores as done here.

Table 3.

	India	Philippines	Singapore	Australia
Rating method	5.66 (100)	5.42 (96)	5.02 (89)	4.91 (87)
'Swiping' method	66.86 (100)	68.17 (102)	66.12 (99)	59.50 (89)

The results show an interesting pattern. In the countries where response biases are expected to play a bigger role (India, Philippines and Singapore), the new method does an excellent job in providing a better base of comparison, as the index range is reduced from 89-100 to 99-102. It should be noted that we do not expect the average for the new method to always be consistent across countries in Asia. Fundamental differences in preference towards the content of the stimuli will continue to cause differences. However, in cases such as this, where there is enough variability in the stimuli to cancel out fundamental differences in preference, we expect the averages to convert to a more comparable level.

In Australia, a country where cultural response biases are less prevalent, no significant adjustment towards a more consistent mean could be observed. We believe that although the new mobile approach helps to level the playing field between countries that exhibit stronger Acquiescent Response Style bias, it is not able to overcome the vast (cultural) difference in response styles between Western and Asian cultures. Taking into account that fundamental differences in preference may cause differences in averages, more cross-country research will be needed to determine if this trend holds true for other Asian and Western markets.

Social desirability

Looking into more detail at the role of Socially Desirable Responses in Asia versus Australia could help further explain the differences between the countries. In tables 4-5, a split has been made between the Asian countries (India, Philippines and Singapore) and Australia. The tables show the average rank of claims, split by the proxy for their social desirability for each of the three methods. For example, the average rank for the three most socially desirable claims in the three Asian countries is 3.2 for rating, 3.4 for MaxDiff and 3.6 for the 'swiping' method.

Table 4.

Asian countries	**Rank for most (#1-3) socially desirable items**	**Rank for average (#4-7) socially desirable items**	**Rank for least (#8-10) socially desirable items**
Rating method	3.2	5.5	7.8
MaxDiff	3.4	5.7	7.3
'Swiping' method	3.6	5.8	7.1

When looking at the data this way, a clear pattern emerges were highly socially desirable claims perform less well in the new method versus the traditional methods and claims that are less socially desirable perform better on average. In other words, the responses of the individual are more likely to deviate from what is socially desirable in the new method. This suggests that the new approach is filtering out some of the response biases and more accurately measures what consumers are really thinking. When looking at the pattern for Australia this becomes even more apparent.

Table 5.

Australia	**Rank for most (#1-3) socially desirable items**	**Rank for average (#4-7) socially desirable items**	**Rank for least (#8-10) socially desirable items**
Rating method	2.0	5.8	8.7
MaxDiff	2.0	6.0	8.3
'Swiping' method	2.0	6.0	8.3

In Australia the average rank across methods is more consistent, with a slight exception for the rating method for less socially desirable items. The more consistent performance of claims across methods suggests that in Australia response bias plays less of a role and the 'swiping' method has less to "correct" for compared to traditional methods. This is in line with expectations that response biases are much stronger in Asian cultures than Western ones. Overall, the results support our hypothesis that traditional methods like rating and MaxDiff favor stimuli that are socially desirable because they rely more on rational processes of the brain. By limiting the need for these rational processes in answering questions, the influence of judgments and social pressure are also reduced.

Engagement

The first two analyses focused on the ability of different methods to accurately predict real consumer preferences. The last analysis focuses on the survey experience. In the chart below the average survey experience per country and method has been mapped out. The scores represent the average of a 5-point rating question with a score of 1 rating the survey experience as 'much worse than other online surveys ever taken' and a score of 5 as 'much better than other online surveys ever taken'. The different lines represent the average scores for the traditional methods versus the new mobile swiping approach across countries.

Figure 1. Engagement scores.

The first thing that stands out is the relatively high scores for the traditional methods. In our traditional client surveys the average engagement score generally falls within the 3.4-3.8 range, so the baseline scores here are higher than usual. This is probably the result of the survey length in our test, which was shorter than it is for typical client surveys. Nonetheless, in three of the countries the new method still raises engagement levels significantly. Existing data from previous studies shows that in actual client projects engagement levels improve even more, with about 0.5 points or more. Overall, consumers tend to appreciate the new way of doing research, which helps keep them engaged and in turn provides us with more meaningful and reliable information.

Conclusion

Although ARS and SDR have long been recognized by both academics and practitioners, the reality is that many of us still use research techniques that are very likely to be negatively impacted by these response biases. For those multinationals aiming to grow their business in Asia where response biases are even more prominent than in western markets, having reliable bias-free insights is critical to their success. This paper aims to start a discussion on how a new generation of research techniques can help overcome these biases and bring us closer to uncovering true consumer behavior. When developing these new approaches it is critical to detach from the past and embrace the world we live in today.

We are now part of a world where mobile devices are at the center of our lives and instant gratification is the norm rather than an aspiration from the past. People have changed and the days that they happily spend half an hour of their day answering a never ending list of questions are over. If we want to really understand the modern consumer we, as a research community, have to embrace the current reality and connect with them in the way they connect with the world around them. This means we have to leverage mobile technologies and use intuitive exercises that are more reflective of the natural ways in which they communicate and express preferences. Swiping, tapping and gamification to keep people engaged and better capture subconscious processes of the brain are the future for anyone who wants to truly understand the consumer of today (and tomorrow).

The Authors

Paul Janssen is Vice President, SKIM, Singapore

Robin De Rooij is Director APAC, SKIM, Singapore

Marina Georgieva is Research Analyst, SKIM, Singapore

.

PART 4
Design Discoveries: The Art and Science of Insight

Slice it, Dice it, and Stitch it Back Together – An Asia Perspective

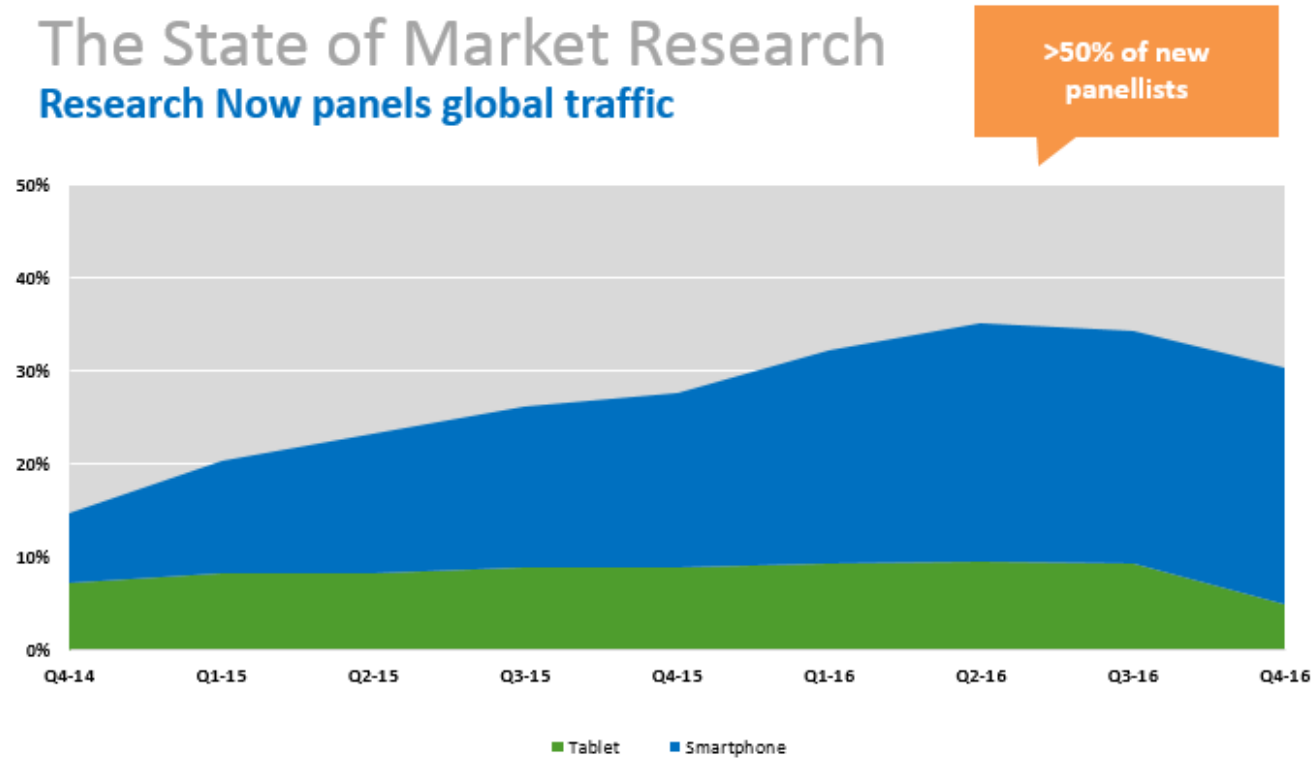
Using split questionnaire design to improve data quality, efficiency and the participant experience

Elizabeth May • Roddy Knowles • Harvir S Bansal • Avik Halder

Introduction

Over recent years, we've seen a steady rise in the number of survey completes being delivered on mobile devices. Today's figures demonstrate the importance of mobile traffic; approximately 40% of online surveys are being taken on a mobile device.[1)] Those signing up to online panels are also expectant of mobile surveys, having already registered on their mobile and when we see that the sign up of new panellists on the Research Now's panels is higher than 50%, it somewhat sets the scene for what to expect in upcoming years. We are at a critical juncture where we must adapt our approach to primary survey research.

Figure 1.



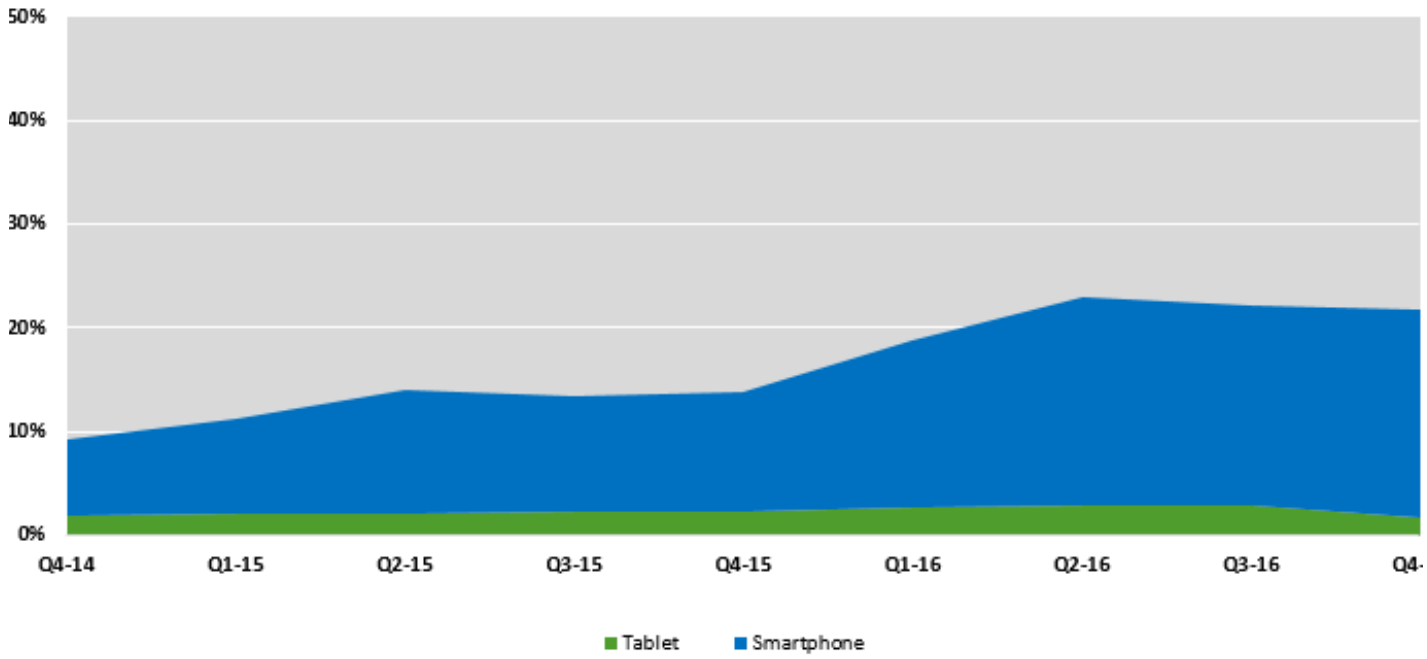
How does this translate for market research in Asia Pacific?

At the end of 2015, 62% of the population in Asia Pacific (2.5 billion individuals) subscribed to mobile services, three-quarters of the region's customer base were in four core markets, that of China, India, Indonesia and Japan. Growth rates in Asia-Pacific are estimated to grow above the global average whereby more than 600 million new subscribers will subscribe to mobile services by 2020.[2)] With this in mind, market research methodologies in Asia-Pacific need to remain and grow with relevance to these consumer technology trends. It is possible that the ability to achieve accurate online survey data across regional locations, not just within metro areas, will become more accessible with the increase of mobile-device adoption.

Combining diminishing attention spans of survey participants and the growth of the mobile-market in Asia-Pacific would not have so much of an impact if the median online survey length in the region was shorter than in other regions across the world. Furthermore, if mobile traffic was similar to that of other regions it would also present itself as less of a barrier to the online market research industry in Asia-Pacific. However, the stats are compelling.

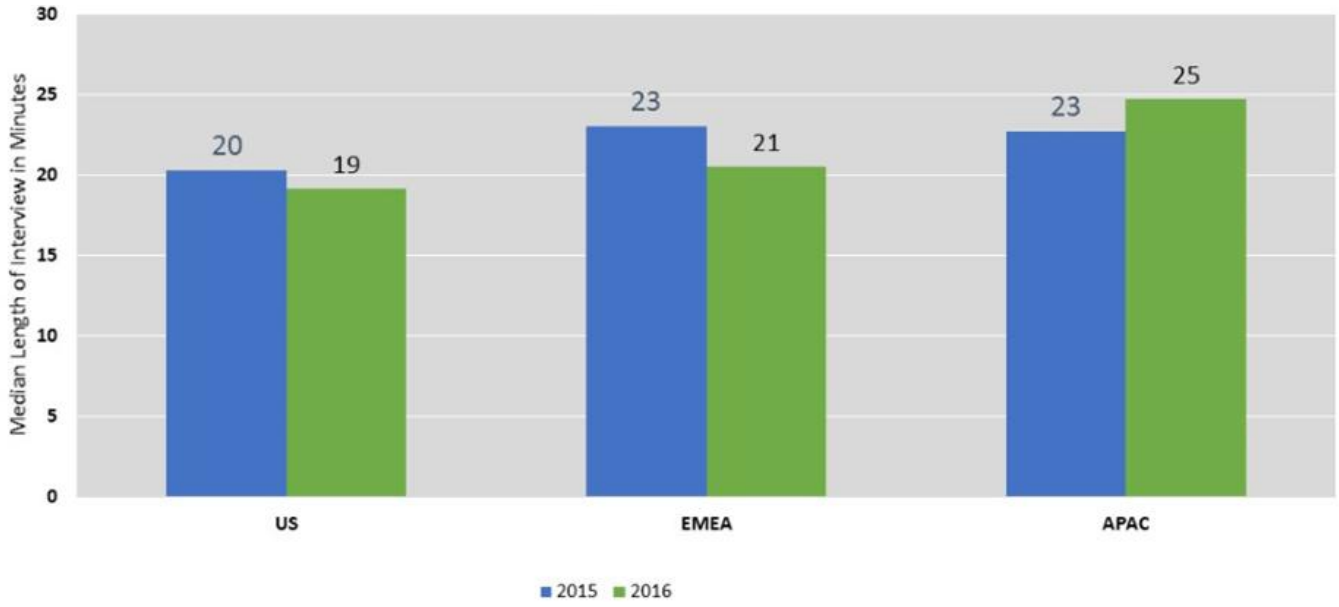
Around 20% of online survey completes in Asia-Pacific are completed on a mobile-device, half that globally at 40%.[3)] The adoption of mobile-friendly survey design in Asia-Pacific is lagging.

Figure 2. Research Now Panels Asia-Pacific Traffic



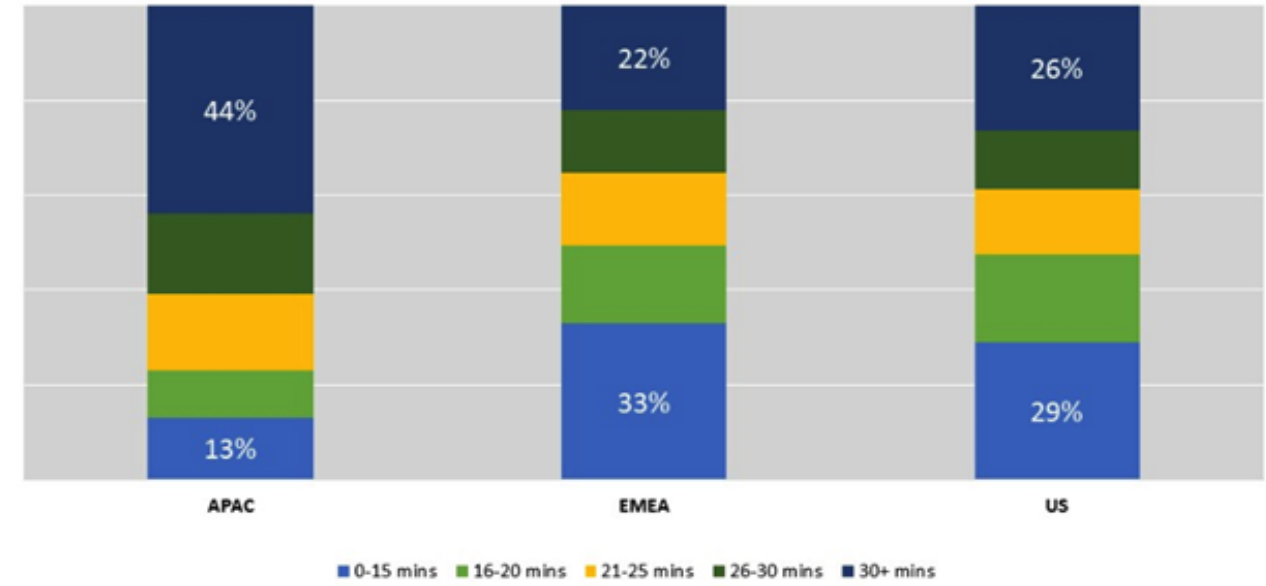
Median survey lengths, in particular for tracker studies, are visibly higher in Asia-Pacific than in other region. Remarkably, trends have only increased in the last year from 2015 to 2016 for this region, whilst the US and EMEA have seen declining interview lengths, starting also from a lower base.[4)]

Figure 3. Online Tracker Surveys – Length of Interview Trend



In 2016, surveys of 30 minutes or more equated to more than 40% of all tracker surveys run in Asia-Pacific markets in the Research Now panels. EMEA and the US online survey lengths are contrastingly lower – to the converse, surveys that are less than 15 minutes in length (and the best candidates to be device agnostic) equate to approximately 30% of all surveys in EMEA and the US.

Figure 4. Online Tracker Surveys – Length of Interview in 2016



What is the way forward for long surveys?

Companies involved in market research carry out large scale surveys for many purposes, mostly with the purpose of informing better business decisions. Researchers are however caught between two forces; delivering the most comprehensive research results for stakeholders and the best research that digs deep in to stakeholders' briefs. In order to achieve this, researchers design surveys that are more detailed and comprehensive, thus requiring lengthier surveys, which is at odds with decreased attentiveness of participants with longer surveys, especially on mobile devices in general and smartphones in particular. As a result, longer surveys' data quality can suffer, propped up by respondent fatigue, such as non-response, early break-off or drop outs and undesired response styles.

One approach to addressing a crucial problem with long surveys is Split Survey Design (SQD) (Vriens, M., Wedel, M. and Sandor, Z., 2001 and Chipperfield, J. O. & Steel, D. G., 2009). SQD allows for surveys to be divided into blocks or "splits," which are in turn distributed to survey participants. Key to survey length reduction is to ask survey participants to provide their input on some, not all, of the splits. Reducing participant burden by using shorter surveys has been shown to improve response rates (Adam and Gale, 1982 and Bean and Roszkowski, 1995). Collins et al (1988) study using telephone surveys explicitly showed that when prospective respondents were informed of the interview length in advance, there was a 14% refusal rate for a 40-minute interview and a 9% refusal rate for a 20-minute interview (Collins et al, 1988). Optimising split design is dictated by minimising the opportunity for information loss without compromising the advantages of completing a comprehensive, but lengthy, survey. SQD, although a resourceful technique to shorten survey length, only generates a partially complete survey response file data set. In order to generate a comprehensive and complete survey data file where various split responses need to be combined, generating a "full" dataset that researchers are comfortable working with. In this work, we describe how to generate complete survey response files using SQD and Missing Value Imputation. The algorithm for SQD used in this work closely follows the work of Adigüzel and Wedel (2005). However, it extends this algorithm by including skip logic, an important component of most current-day survey designs.

Though the idea of creating split surveys is not new, selecting an optimal design strategy is still an active area of research. One of the key challenges in successfully executing SQD, and similar approaches to survey modularisation and data imputation, has been the time and effort involved in the often highly manual front-end design and back-end imputation processes. This paper outlines a highly automated process for implementing SQD – and one that can be implemented with surveys that have not been specifically designed with modularisation in mind. Another usefulness of split questionnaire design is that survey costs can be reduced. Shorter surveys mean lower sample costs. Thus, SQD produces a more cost-effective approach to deliver short and efficient surveys.

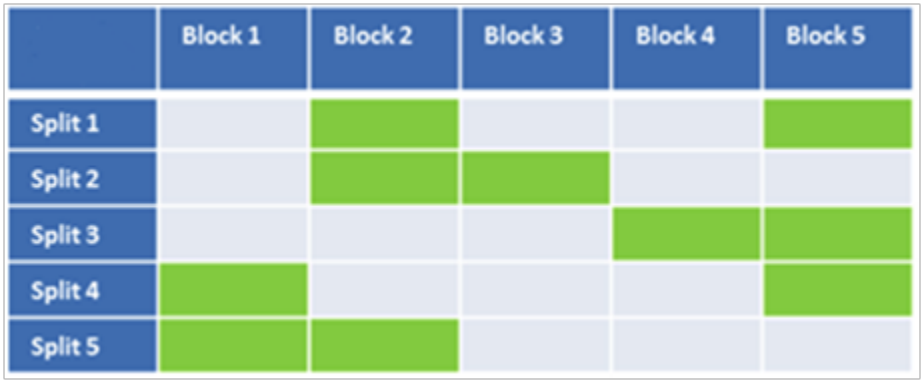
Split survey design

As previously mentioned, the express purpose of SQD is to break long surveys into multiple, shorter, "split" surveys thereby reducing the time and effort it takes for participants to complete. It has been demonstrated that reduced participant burden improves data quality when SQD is implemented (Adigüzel and Wedel, 2005). The challenge, however, is to accommodate the questions in a longer survey so that the objective of data comprehensiveness is balanced with the objective of reduced participant burden. The study highlighted here, as is common with many surveys, contains binary (e.g. yes or no) and discrete (e.g. scale rating) questions. Additionally, surveys may contain questions that require providing continuous data with decimal points to questions such as currency based questions, or in countries that use the metric system for height and weight.

In order to implement a general SQD algorithm, the questions that are closely related thematically are placed into blocks (Raghunathan and Grizzle, 1995). Let's define two terms - "block" and "split." A *block* is a group of questions/variables in the survey that always appear together. For example, if there are 20 questions, and five blocks in the design, then each block could contain four questions. However, blocks may not always be uniform, and may vary depending on the type of questions grouped together. The algorithm can decide what should be the optimal number of blocks. On the other hand, a *split* is a group of respondents who view the same distribution of blocks. For example, if split 1 in the design receives blocks 1 and 2, it means that all the respondents grouped together in that particular split will see the questions from blocks 1 and 2 only.

An example of a 5-block, 5-split design is illustrated in figure 5.

Figure 5. A 5-block 5-split design example



Integrating skip logic in SQDI

Most surveys contain filtering questions that flow respondents across the survey as a result of the responses that they provided earlier on in the survey. For example, "Which OS does your primary mobile phone run? 1. Android, 2. iOS or 3. Other." If "Android" is selected, then the next question will be related to the Android OS; conversely, the question would be related to iOS if it were selected. The split survey design algorithm discussed in this paper can be adapted to incorporate skip logic. The appendix to this paper provides examples of the most commonly used forms of skip logic.

Asia-Pacific case studies – Both a tracker and adhoc survey

The SQD process was implemented on two specific case studies in Australia and Japan, both related to sports. The first study was a long-established and complex tracker survey of 32 minutes, based on general sport participation across Australia. The second survey was an adhoc survey that had been previously run to better understand the Japanese behaviour of supporters during football games, more specifically, which football players they like, how much they engage with social media during the games and perform other interactions on the Internet.

In both cases, the optimal split survey designs were remarkably different. Data from the prior study was utilised to establish the optimal split survey design. For the tracker, three months of data was analysed and just the one wave of data was available for the adhoc survey.

In the case of the tracker survey, a 40-split dynamic-block design across n=1,092 participants was implemented. For the adhoc Japanese survey, a 4-split 4-block design was created for 301 participants for the SQD version. Both of the modularised versions were conducted alongside the full survey in which the entire, un-modularised survey was used.

Let's take a deeper look at the Australian study as an example of how the methodology was applied?

The survey was designed with several levels of skip logic hierarchy, requiring that SQD was performed on every level. As the survey was highly complex, the split-block distribution was performed dynamically on each level. On the first level of skip logic hierarchy, a 10×10 design was created. On the following levels, 6×6, 4×4 and 2×2 designs were applied. This resulted in a total of 40 different splits. Figure 5 provides two examples, a 10x10 matrix and a 6x6 matrix.

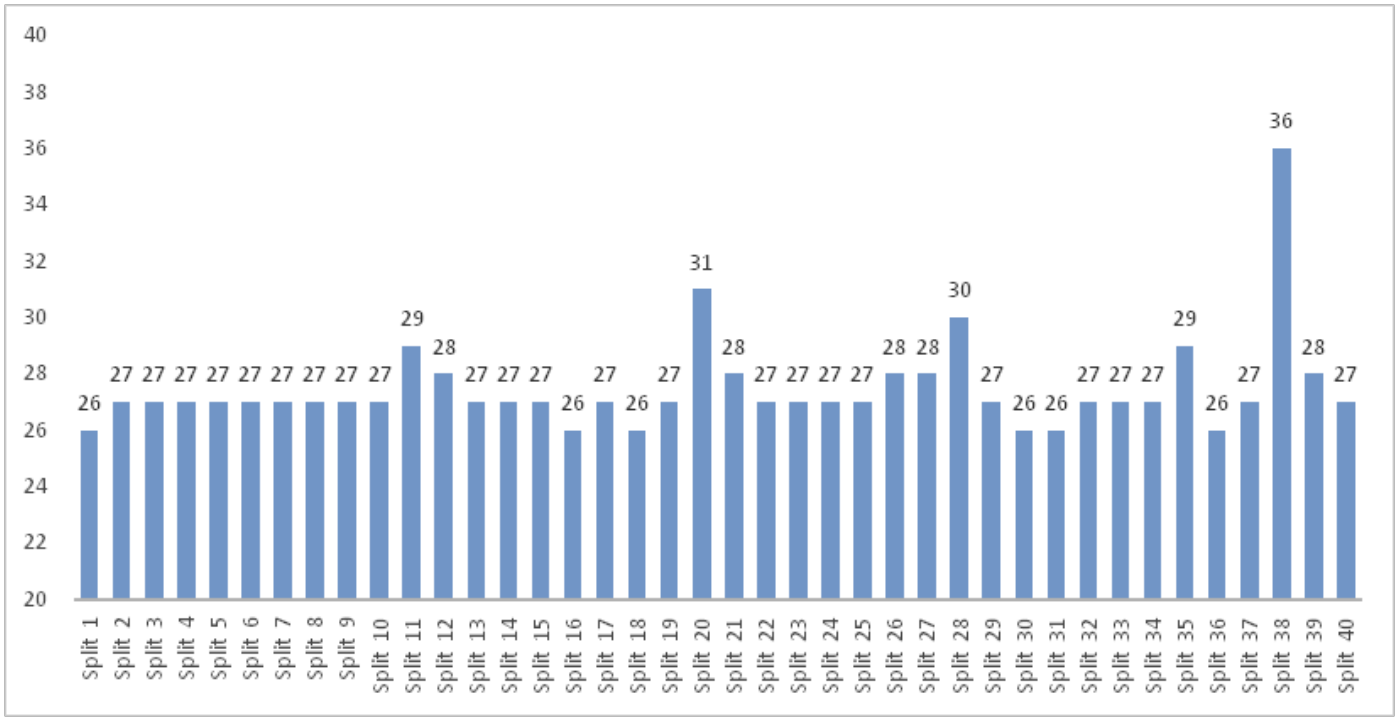
Figure 6. Split survey design, where blocks marked green are provided in the-split

Level 1 Design	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	Block 7	Block 8	Block 9	Block 10
Split 1	■			■		■	■			
Split 2		■	■		■			■		
Split 3						■	■	■		■
Split 4			■	■	■				■	
Split 5	■	■		■			■			
Split 6	■		■		■				■	
Split 7		■				■		■		■
Split 8			■			■			■	■
Split 9	■			■				■	■	
Split 10		■			■		■			■

Level 2 Design	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6
Split 1		■		■	■	
Split 2			■	■		■
Split 3	■	■			■	
Split 4		■	■			■
Split 5	■			■		■
Split 6	■		■		■	

812 complete responses were collected from the full survey, while 1,092 completes were collected for the SQD testing survey. The sample sizes for each split are shown in figure 7.

Figure 7. Population of participants per split



The missing data due to incompleteness of split design was imputed using the Monte Carlo imputation. The results presented below compare complete survey and SQD data after imputation. Two types of comparisons conducted on an aggregate level are:

1. *Visual similarity*: Comparison of frequency plots of each possible answer of each questions
2. *Statistical similarity*: Mathematical judgment whether data from full and imputed surveys have similar distribution

The comparison was done for all survey completes, as a total, and then for each device type so the analysis could demonstrate the effectiveness of the SQD algorithm used in generating quality data.

Figures 8 and 9 displays the overall visual similarity. The stacked bar charts show the proportion of participants from both surveys for all possible answers of a question. The plotted participant ratio is the percentage ratio of the numbers of people in both surveys who responded with a particular answer for the question. Blue represents full survey data and orange represents SQD imputed data. If both bars in each stack are 50%, then they are exactly similar. Here, we can see very similar ratios for both binary and categorical questions.

Figure 8. Two different binary questions: blue denotes full survey data and orange denotes SQD imputed data

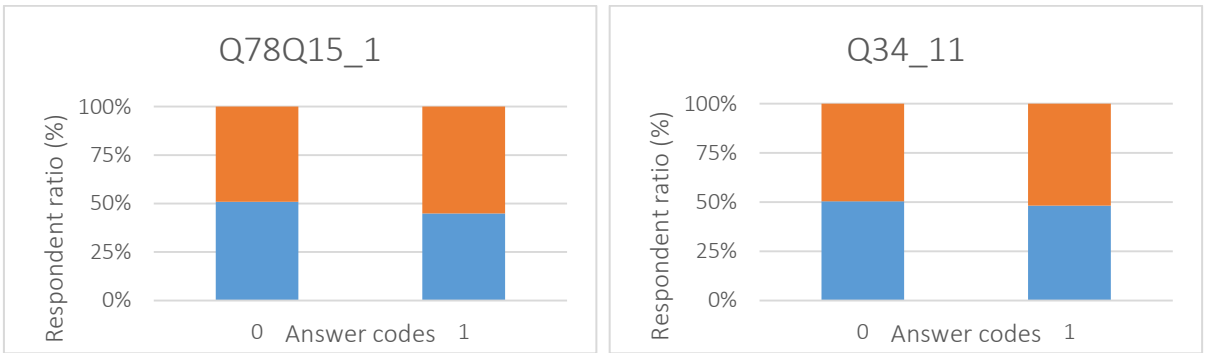


Figure 9. Two different categorical questions: blue denotes full survey data and orange denotes SQD imputed data

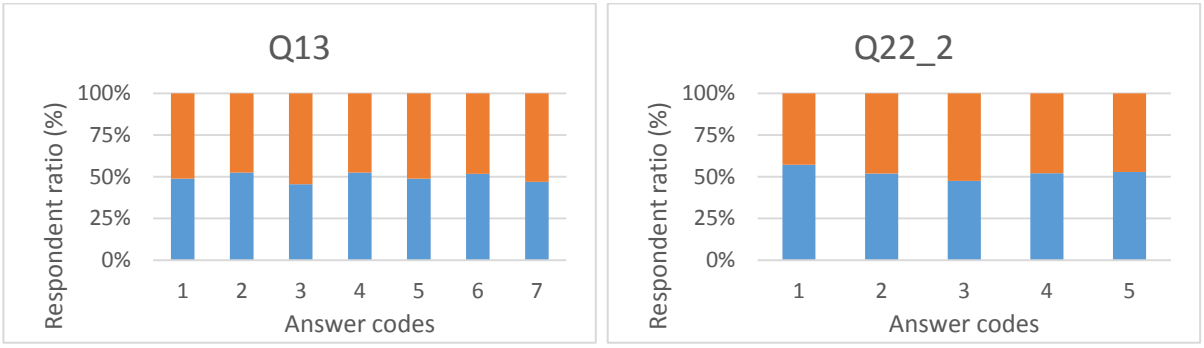


Table 1 displays overall statistical accuracy. The goal of the statistical comparison was to verify whether the imputed and the full survey data could be treated as two samples from the same population, i.e. they followed the same distribution. Chi-square test was used for the categorical questions, and Z-test for the binary questions. The accuracy rate is defined as the ratio between the number of questions for which the imputed and the real data are similar. A significance level of 98% was used for calculating the accuracy rates. As seen below, the data delivers a very high rate of accuracy at both binary and categorical questions. The tests verified that the distribution of the imputed responses is near identical to the distribution observed as the same question in the full survey. This similarity of statistical distributions speaks to the robust and sound capability of Splice™.

Table 1. Statistical comparison of full survey vs. imputed data for all devices at 98% level of significance

	Accuracy Rate
Binary questions	91%
Categorical questions	92%

The data resulting from those who completed the survey on a smartphone delivered a higher level of statistical accuracy as below.

Table 2. Statistical comparison of full survey vs. imputed data for mobile devices at 98% level of significance

	Accuracy Rate
Binary questions	93%
Categorical questions	95%

We further assessed improvements to qualitative measures of survey data quality using SQD, as compared to the full survey as shown in table 3. The table shares data for the poor verbatim rates (ratio of number of poor verbatims as a percent of the total number of entries), drop-off rates (ratio of people who dropped-off after seeing the first question) and the median time to complete the survey.

Table 3. Statistical comparison of full survey vs. imputed data for mobile devices at 98% level of significance

	Full survey data	Splice™ survey data
Poor verbatim rates	2%	1%
Drop-off rates	17%	19%
Median Time to Complete	32 minutes	25 minutes

The Splice™ SQD version achieved significant improvements to reduce the length of interview from 32 minutes to 25 minutes. For the Splice™ guided survey, the drop-off rate did not drop as would have been desired. It can be inferred that a survey of 25 minutes, although 22% shorter than its original, is still considered too long to be mobile-friendly.

It must be kept in mind that even when splitting surveys, there are always some fixed overheads. The screener and demographic questions must be asked to all respondents, which are then used for imputation purpose. Also, there are questions which the SQD algorithm decides to ask all respondents due to skip logic rules.

How did the Japanese adhoc survey perform as a Splice™ SQD version?

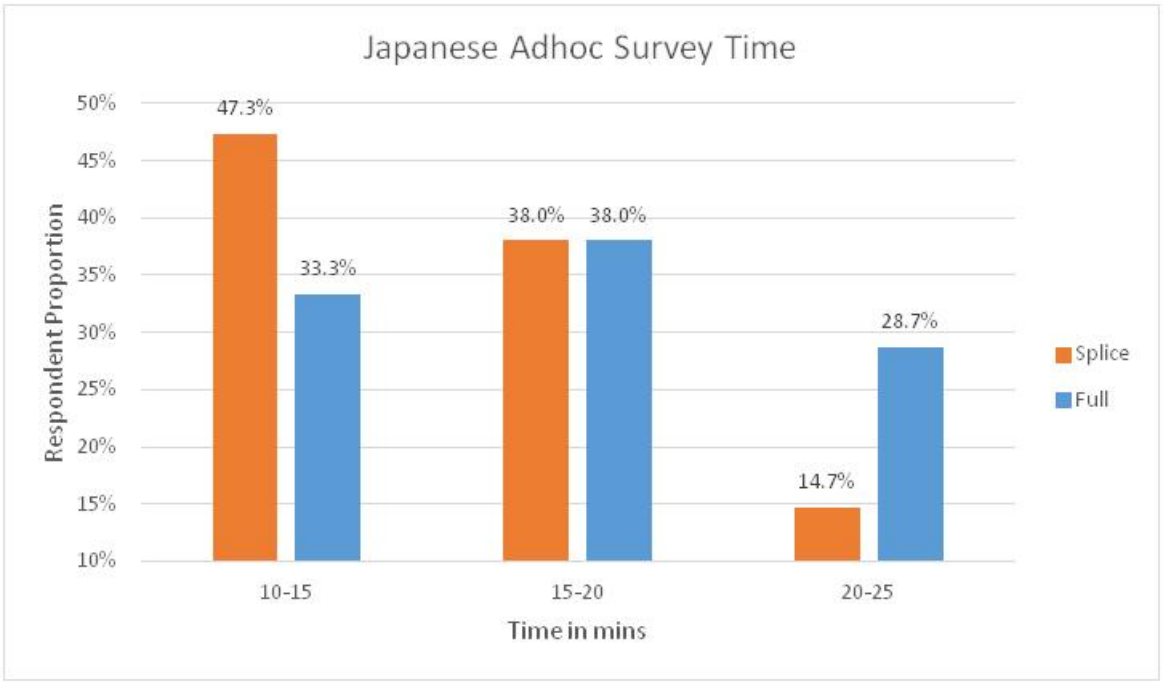
Key measures relating to Poor verbatims and Drop-off rates were reduced for the Splice™ survey version, indicating a more acceptable survey design was delivered under this methodology.

Table 4. Improvement of survey data quality using SQAD for an adhoc survey in Japan

	Full survey data	Splice™ survey data
Poor verbatim rates	2.5%	6.1%
Drop-off rates	6.5%	3.3%

It is important that interview length is improved across the board, for those who take longer or shorter time to complete a survey. As a result, we analysed the distribution of the time required to complete the survey across all length of interview groups for survey participants. The distributions are shown in figure 9 for the Japanese adhoc study.

Figure 10. The distribution of time to complete both surveys. The distribution is skewed to the left for Splice™ survey, which is desirable

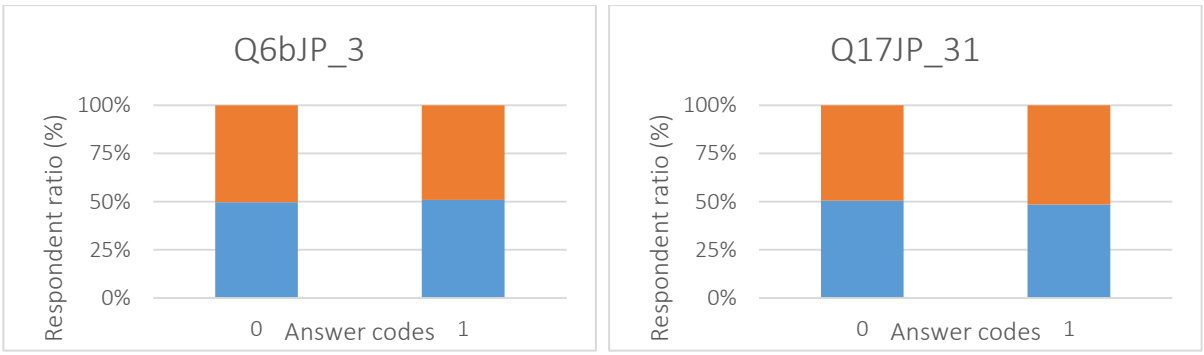


For the regular un-modularised survey, 33% of respondents completed the survey in 10-15 minutes whereas 29% of respondents needed 20-25 minutes to complete the survey. In an ideal world, it would be preferable to improve the distribution of time to complete the survey in two ranges; those who completed the survey at the most favourable and shorter timeframe would increase in number and the percentage of those completing the survey in excess of 20 minutes would reduce. A "see-saw" improvement effect was achieved – 47% of the Splice™ version participants completed the survey in 10-15 minutes and only 15% of participants completed the survey in 20-25 minutes. This verifies the effectiveness of Splice™ to decrease the time needed to complete the survey.

Does the Splice™ SQD version work as well on Mobile Devices as Desktop?

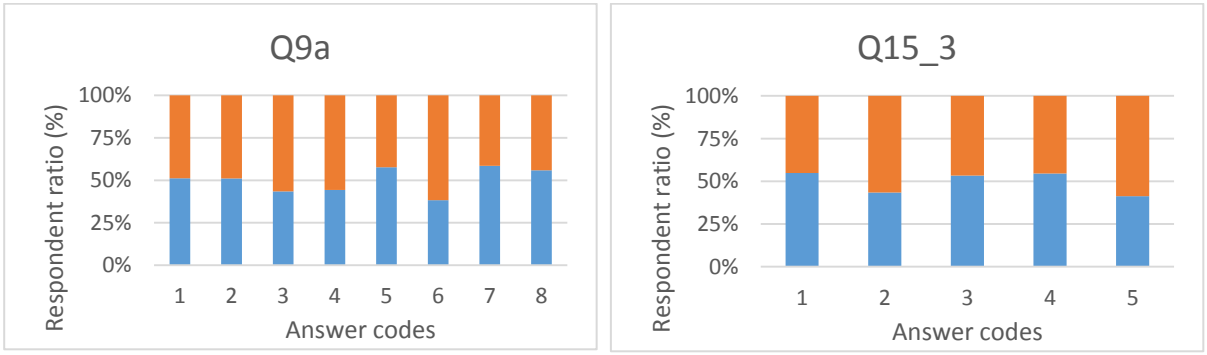
As mentioned above, quantitative comparisons were also performed at the device level to assess the device specific effectiveness of SQD surveys. In this case, we provide the data pertaining to the Japanese adhoc survey. Figure 11 shows the visual similarity using frequency plots for Desktop devices for binary questions and Figure 12 shows the comparison for categorical questions. Figures 13 and 14 show the same for mobile devices.

Figure 11. Bar charts showing visual similarity for binary questions for desktops.



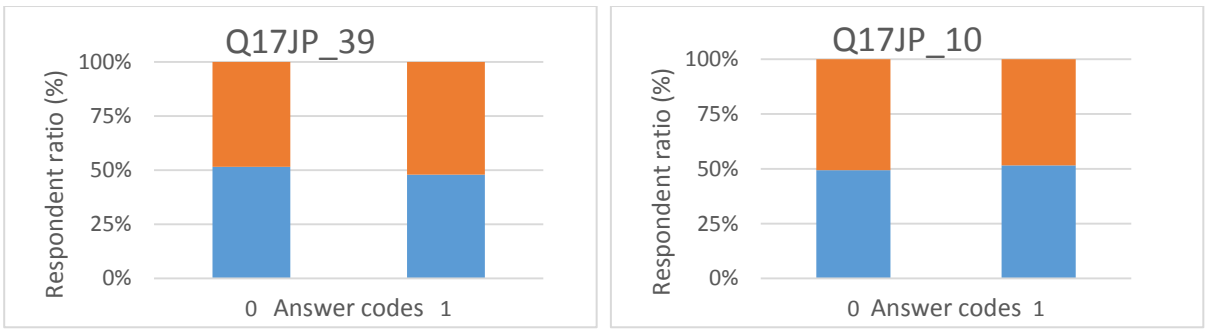
Blue denotes full survey data and orange denotes SQD imputed data

Figure 12. Bar charts showing visual similarity for categorical questions for desktops



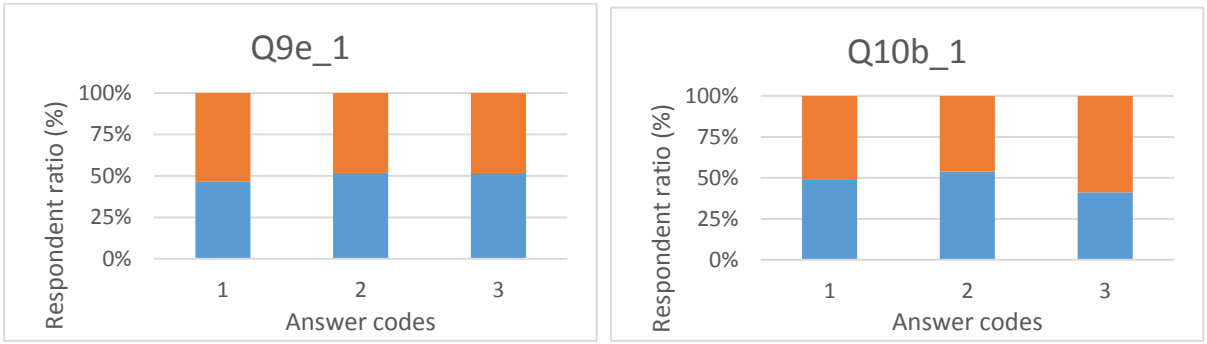
Blue denotes full survey data and orange denotes SQD imputed data

Figure 13. Bar charts showing visual similarity for binary questions for mobile devices



Blue denotes full survey data and orange denotes SQD imputed data

Figure 14. Bar charts showing visual similarity for categorical questions for mobile devices



Blue denotes full survey data and orange denotes SQD imputed data

Table 5 shows the statistical comparison for the Japanese adhoc survey between real and imputed data for specific devices. Whilst the adhoc Japanese study showed a higher data accuracy versus the Australian tracker study, both delivered a high success rate. Data was near identical across the full survey and the Splice^{TM} versions.

Table 5. Statistical comparison of real vs. imputed data for desktop and mobile devices at 98% Level of Significance

	Accuracy Rate Desktop	Accuracy Rate Mobile Device
Binary questions	90%	94%
Categorical questions	96%	96%

As seen in table 5, we can see very high accuracy for mobile devices, especially on the categorical questions. It can be also observed that the device-specific accuracy rates are higher than the overall accuracy rate.

What was the effect of Splice™ SQD on panellist experience?

From a qualitative perspective, survey participant experience is both desirable and necessary to achieve sound data for analysis. Key to demonstrating that the Splice™ technique to modularisation is appropriate on this level is to see if the panellist experience was improved. We observe this through quantitative and fieldwork measures however we also asked participants to rate the survey once completed. Direct panellist feedback is important.

The results were very positive, in particular for the mobile version of the Splice™ survey version. Table 6 demonstrates the effectiveness to drive panellist engagement.

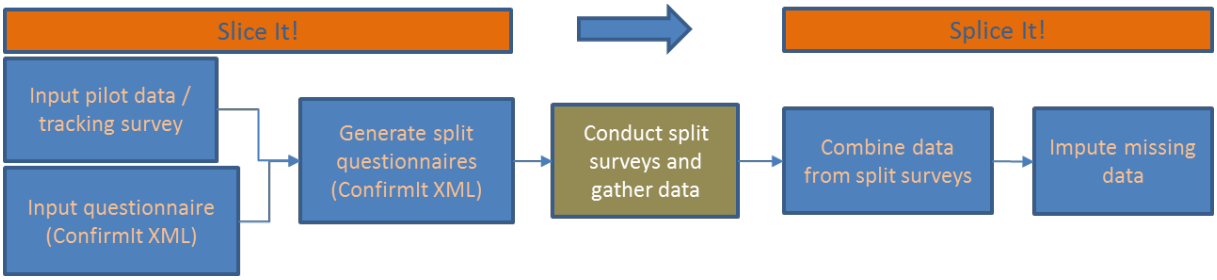
Table 6. Participant feedback on survey satisfaction – Australia Tracker Survey

Agree/Strongly Agree	Full Survey	Splice™ Survey (Overall)	Splice™ Survey (Mobile)
Survey was well written	45%	**45%**	**64%**
Allowed to express opinion	42%	**43%**	**69%**
Time and effort required	38%	**39%**	**44%**
Overall experience	41%	**43%**	**67%**

Splice™ software

The algorithm for SQD with skip logic implemented and missing data imputation is encapsulated in the SPLICE™ software. Figure 14 demonstrates the project flow that the software applies itself to. The Slicing part of the software takes the full survey and pilot/previous wave tracking survey data as inputs and generates split surveys as outputs. In the Splicing phase, the data from the split surveys are combined and the missing responses are imputed. The software uses languages C++ and R. The main software is written in C++.

Figure 15. How Splice™ works



The software integrates easily with ConfirmIt whereby the ConfirmIt platform provides an interface for program surveys. This platform is available on-demand as Software as a Service (SaaS) and surveys can also be hosted on the ConfirmIt server itself.

The SPLICE v1.0.1 supports ConfirmIt version 18.5. The interface is simple, user-friendly, and mobile-friendly. The interface is designed using HTML5 and CGI C++ so that it can run C++ coded software in the background.

Conclusion

Although nothing new, survey modularisation offers an attractive solution to comprehensive data collection, whilst mitigating survey participant disengagement. This is particularly true for the Asia-Pacific region. In the context of changing consumer digital behaviour, and an increasing reliance on smart-devices as consumers' primary connected screen, long surveys are a fundamental challenge for the survey industry. Furthermore, for any solution to be widely adopted it needs to be consistently easy to deploy and deliver comparable data quality.

This work has built on previous modularisation work, using a robust methodology to demonstrate a strong and tested approach to alleviate respondent burden. Our approach draws on computational automation techniques to optimise the process of modularisation. In the two studies, we used automated statistical interpretation of survey behaviour to devise the best survey design strategy for modular data collection.

Further research and development must address the need for one-touch deployment of any modularisation solution. This approach is a crucial evolutionary step towards the machine learning that will ultimately allow SQD to be implemented 'on-the-fly' for a truly dynamic survey which responds in real-time to the current respondent.

Endnotes

1. Research Now Group Inc. global survey stats 2014-2016
2. The Mobile Economy Asia Pacific 2016 – The GSMA (www.gsma.com/mobileeconomy/asiapacific/)
3. Research Now Group Inc. global and regional survey stats 2014-2016
4. Research Now Group Inc. global and regional survey stats 2014-2016

References

Vriens, M., Wedel, M. and Sandor, Z. (2001), "Split-survey designs: A new tool in survey design and panel management", *Marketing Research*, Vol. 13, No. 2, pp. 14-19.

Chipperfield, J. O. & Steel, D. G. (2009), "Design and estimation for split survey surveys" *Journal of Official Statistics: an international quarterly*, Vol. 25, No. 2, pp. 227-244.

Adams, La Mar L. and Darwin Gale (1982), "Solving the Quandary between Survey Length and Response Rate in Educational Research", *Research in Higher Education*, Vol. 17, No. 3, pp. 231-240.

Bean, Andrew G. and Michael J. Roszkowski (1995), "The Long and Short of It: When Does Survey Length Affect Response Rate", *Marketing Research*, Vol. 7, No. 1, pp. 21-26.

Collins, M., Sykes, W., P. Wilson, and N. Blackshaw (1988), "Nonresponse: The UK Experience," In: Groves et al. *In Telephone Survey Methodology*, Wiley and Sons, New York, NY, pp. 247-256.

Feray Adigüzel and Michel Wedel (2005), "Split Survey Design", *Ross School of Business Paper,* No. 966.

Feray Adigüzel and Michel Wedel (2008), "Split-survey Design for Massive Surveys", *Journal of Marketing Research*, Vol. XLV, pp. 608-617.

Raghunathan, Trivellore E. and James Grizzle (1995), "A Split Survey Design", *Journal of American Statistical Association*, Vol. 90, No. 429, pp. 54-63.

Johnson, J., Siluk, L., & Tarraf, S (2014), "Techniques for fusing survey modules.", *International Journal of Market Research Vol. 56 Issue 4, pp. 425-442*.

The Authors

Elizabeth May is Vice President Asia-Pacific Client Development, Research Now, Australia

Roddy Knowles is Director, Product and Research Methodology, Research Now, United States.

Harvir Bansal is CRO and EVP Analytics, B3Intelligence, Canada

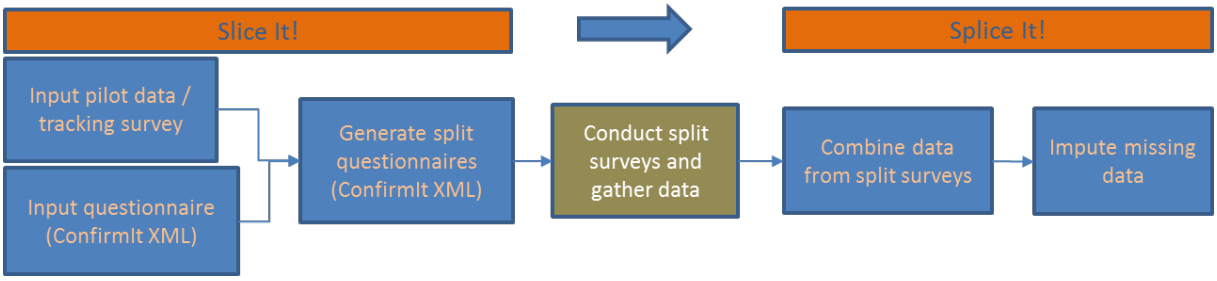
Avik Halder is Manager, Data Science,, B3Intelligence, Canada

Appendix – Examples of skip logic

Case 1: Survey with a main skip logic question with many or a few dependents

There exist a large number of dependent questions corresponding to one skip logic question. For example, a survey contains the question: “From which of the following companies have you made a purchase of consumer electronics, appliances, or entertainment products such as music or movies in the past 30 days?” Based on the selected answer, a participant is categorized as “purchaser” or “non-purchaser.” The questions that follow are marked to be asked to one of these specific categories. For example, [ASK IF PURCHASER] “Why did you decide to buy your product(s) from the company?” The label “[ASK IF PURCHASER]” means this question is dependent upon the previous skip logic question. The skip logic model is purposely designed so the survey is dependent on the [ASK IF PURCHASER] variable where a large number of dependent questions exist. The alternative is that only one or two dependent questions exist for a skip logic question. The solution of this scenario is to implement SQD at each level depending on a threshold determined by the number of dependent questions. Figure 16 illustrates the outcome of selecting few dependent questions:

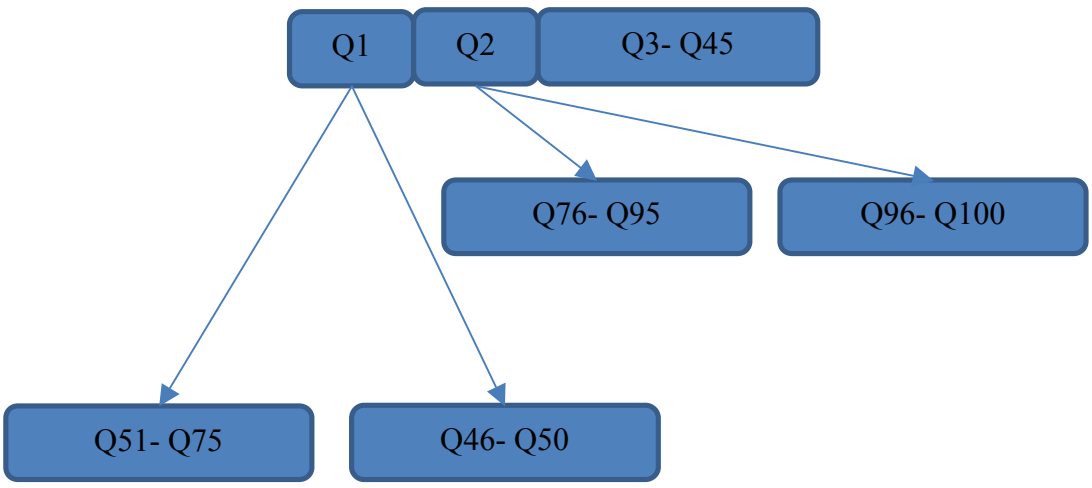
Figure 16. Survey with one skip logic question with many or a few dependents



Case 2: Survey with a multiple skip logic questions – when more than one skip logic questions exist and each one has a number of dependent questions

For example, a survey comprised of 36 questions may have Q6-Q16 dependent on Q1 and Q32-Q33 depend on Q31. Q1 and Q31 are independent of each other. The solution in this scenario is to initially isolate the skip logic questions, distribute them equally to participants and perform SQD with dependents. Figure 17 demonstrates the scenario. SQD is performed on the left branch of Q1 and the left branch of Q2.

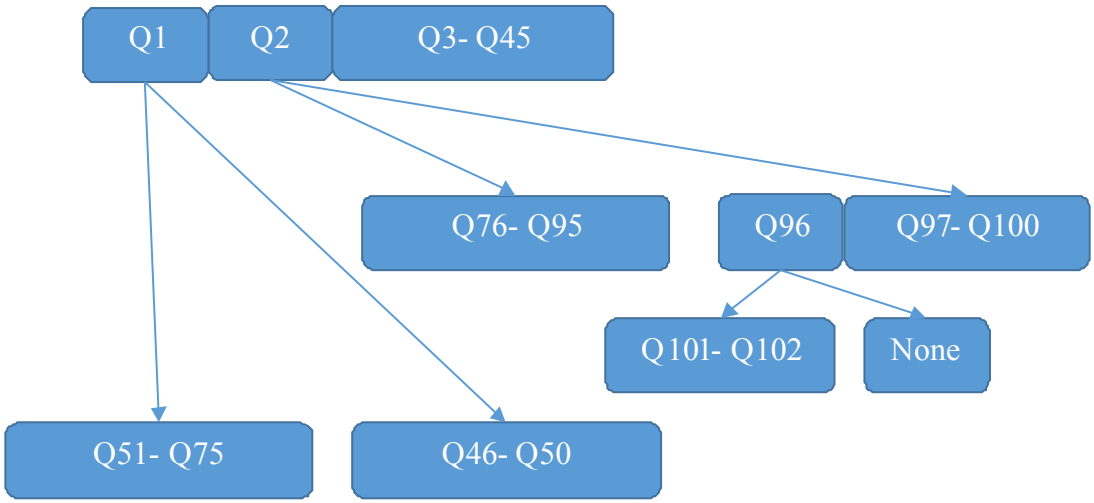
Figure 17. Survey with one skip logic question with many or a few dependents



Case 3: Survey with skip logic within dependents

If skip logic exists within dependent questions, a tree-like hierarchical structure can be found in the survey. Figure 18 shows Q96-Q100 are dependent on Q2 and follows its right branch. Again, Q101-102 follows the left branch of Q96. In this case, one should follow previous guideline for each level in the hierarchy of the survey. The algorithm will process the lowest level first and systematically filter the upper levels until it reaches the top of the hierarchy. If there are multiple skip logic questions on any level, they are handled using the same approach illustrated in Case 2.

Figure 18. Survey with skip logic within dependents



Case 4: Survey with skip logic questions with termination branches

One branch of the skip logic has a “terminate” instruction. This type of question is normally inserted at the beginning of a survey where the answer determines whether a person is a qualifying candidate for the survey. These types of questions are necessary and irreplaceable because they screen the viable participants for the survey. SQD is not performed on these types of questions because they are only meant to eliminate non-participants. SQD is executed on surveys being answered by viable participants who pass the screening process. However, there may be a second set of survey questions with a screener question. In this situation, if the participant chooses the termination branch of the first screener, the second screener will be asked, and if the participant answers with a viable response, the second survey continues and SQD is applied to the second survey applying the previous guidelines. If the participant answer is not viable, the second screener will lead to the termination of the survey. If there are no more alternative surveys, the whole program will terminate.

The Rules of the Game

A cross cultural study of gamification techniques in Asia

Jennifer Serrano • Pete Cape

Introduction

On September 29th, 2007, the first episode of *Bǎi Wàn Zhì Duō Xīng* was shown on the Chinese TV station GhuiZhou TV. The show features a quiz competition in which contestants attempt to win 1,000,000 Chinese Yuan by answering a series of questions. This Chinese game show was based on the original British program; "*Who Wants to Be a Millionaire?"* Versions were also shown in Japan, Australia and in over 100 other territories around the globe. The game show, according to Television Business International, is the most popular TV contest and the runaway winner in terms of audiences across the world.[1)]

What makes it so engaging to people across different cultures?

Dan Ariely, professor of Psychology and Behavioral Economics at Duke University, says in his TEDxMidwest talk[2)] that what gets us to care, act and do things is far from many people's simplistic view of motivation - which is to enjoy leisure. He argues that "What truly motivates us is not relaxation, it's not comfort, it's other things. It's about achievement, it's about conquering and it's about pursuing some goals."

Unlike some older game shows which were designed to make the contestants feel at ease, the set, along with all the other elements of the *Millionaire* game show, are presented in such a way as to make the contestant feel uncomfortable. For example, the musical score was created to increase tension as the contestant progresses through the game. The lighting system was programmed to make the contestant feel as if they are outside a prison when an escape is in progress.[3)]

How exactly might that work?

Gabe Zichermann, proponent of game mechanics in business and other non-entertainment platforms, posits that the power of games is tied to the concept of intrinsic reinforcement. He says that "anytime you *challenge* yourself to something, anything at all, doesn't matter how big or how small and you *achieve* that thing, your brain secrets a magical little bit of neuro transmitter called Dopamine."[4)] Neuroscientist Ian Robertson, in his study of how power affects the brain, describes Dopamine as the 'medium of success.' He further explains that "Dopamine has important effects in the brain particularly in the functioning of the frontal lobe which is critical for motivation, for focus and for setting goals." [5)]

This perhaps explains why people like playing games.

A game, however, is only as good as its rules. How well the game is played is defined by how well the rules are followed. In our world of market research we have learnt that by introducing arbitrary rules, the question becomes *more challenging*. A clear set of rules and goals also communicates to participants what they should achieve and *how it can be achieved* (Puleston and Sleep, 2008).[6)]

However, what is unclear is whether all people, in different countries and cultures, react in the same way to a set of stimuli or rules.

Hofstede's cultural dimensions theory suggests there are systematic differences in national cultures based on four primary dimensions: individualism-collectivism; uncertainty avoidance; power distance (strength of social hierarchy), and masculinity-femininity (task orientation versus person orientation).

Uncertainly avoidance, for example, defines society's tolerance for ambiguity or unstructured situations. Societies that score a high degree in this index opt for stiff codes of behavior or guidelines as they

generally rely on absolute truth. Societies with a lower degree, on the other hand tend to impose fewer regulations as the people are more accustomed to ambiguity and differing thoughts and ideas. We would expect cultures high on Uncertainty avoidance to be more likely to adhere strictly to rules.

The second dimension that we look at is whether a society is collectivist with strong social bonds or individualistic. Collectivist countries are characterized by harmony, avoidance of confrontations and *more conformity behavior*. In contrast, individualist countries are characterized by an acceptance of confrontations and *lower emphasis on conformity* and harmony. In a collectivist society, when rules are set, we would expect people to conform to them. In our study, this means they should exhibit less satisficing behavior when answering the standard question given the implied rule: "write down as many as you can." Whereas in individualistic cultures, being focused on "I," people ought to be more motivated by games and the possibility of winning.

By linking to this established cultural dimension theory, particularly looking at these two dimensions, we will explore how the differing cultural tendencies might affect someone's willingness to abide by and play by the rules.

Setting the rules and getting everyone to agree on the rules is important. The more we understand how the effects of society's culture and values relate to behavior, the more successful we will be in adapting gamification techniques at a local level and thereby unlock its promise of improving participant engagement and providing richer data.

In this paper we show key findings from new research that provides insights on the practical application of gamification techniques, particularly in APAC. Specifically we ask:

1. Does everybody play games?
2. Do all cultures play games to the same extent?
3. What is a good rule to set?

Methodology

To investigate the effect of cultural differences in executing a gamified question, we conducted an online study across five countries in the APAC region. Using Hofstede's country levels score, the following countries were chosen to reflect the variation observed on two of Hofstede's dimensions:

China	(Collectivist and tolerating Uncertainty)
Hong Kong	(Collectivist and highly tolerating of Uncertainty)
Singapore	(Collectivist and extremely tolerating of Uncertainty)
Japan	(extremely high Uncertainty Avoidance)
Australia	(extremely Individualistic)

Our experiment involved a sample of 3,000 participants with 600 interviews per country. Our data set consists of two distinct experiments:

The first experiment looked at three versions of a question on spontaneous brand awareness within the Mobile Phone category.

- Cell 1: 200 adults answered the standard question ("write down all you can think of")
- Cell 2: 200 adults answered the gamified question using rule 1 ("maximum of 10 brands")
- Cell 3: 200 adults answered the gamified question using rule 2 ("at least 5 brands")

In the second experiment, we asked them to describe their favorite brand using three different sets of rules:

- Cell 1: 200 adults answered the gamified question using rule 1 (use exactly 5 words/characters)
- Cell 2: 200 adults answered the gamified question using rule 2 (use exactly 7 words/ characters)
- Cell 3: 200 adults answered the gamified question using rule 3 (use exactly 10 words/ characters)

Does everybody play games?

In his article published in *Harvard Business Review*[7], Nick Shore, Senior Vice President of Strategic Consumer Insights and Research at MTV, writes about how their study on Millennials' digital habits confirmed their intuition that a "game-like metaphor" applies to almost every aspect of a Millennials' life. He further highlighted how this generation believes that "life can be less stimulating than gaming." No wonder then that many industries are finding ways to employ gamification, or "the use of game design elements in non-game contexts" (Deterding, Dixon, Khaled, & Nacke, 2011[8]), to engage their customers, especially the younger generation.

Gamification in online surveys is used to describe a wide variety of techniques aimed at increasing participant engagement (for an overview see Keusch and Zhang, 2015[9]). Suggested approaches from Puleston and Sleep (2011[10]) include rephrasing questions to sound more game like, or making the way questions are answered more game like. On many occasions, it has been shown to improve user engagement and, importantly, to improve data quality. Gamification has been especially helpful in improving response quality to open questions. It works primarily because it plays into one of the core drivers of motivation ─ the need to achieve and to feel competent at the task. By turning the question into a game we put in place rules that allow participants to focus on the task and to win at research.

But who actually does "play" games?

In order to examine this we began by replacing the standard spontaneous brand awareness question with a "little game." In the standard (non-gamified) question we ask them: "Please write down the name of all the mobile brands you can think of." To turn this into a game we added some game-like elements. We set the objective and the criteria for winning (naming ten brands) and the rules (60 seconds to write them all down). If people are playing the game, we should expect that they will give us more brands in our gamified version. Remarkably, the results show (in table 1) that the average number of brands mentioned increased in the gamified question across all of the five countries! In the standard version, participants name, on average, two to five brands. This goes up to four to seven brands, as high as an 80% increase, in the gamified version.

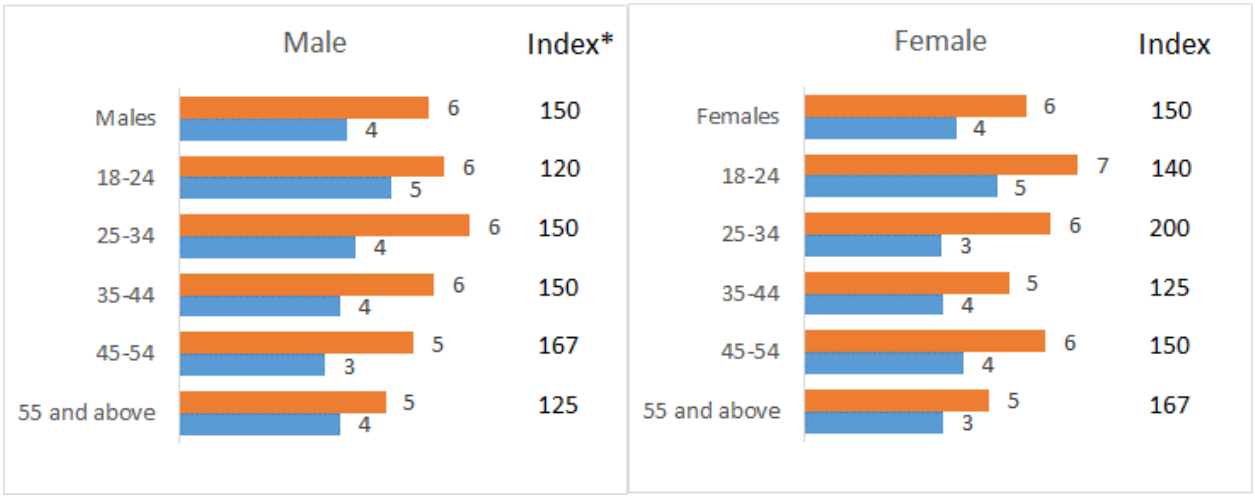
Table 1. Average number of brands by country

Country	Standard	Gamified (max of 10)
China	5	7
Singapore	5	7
Australia	4	5
Hong Kong	3	6
Japan	2	4

It is not only young males (who are often thought to be more predisposed to playing games) who play. People of all ages, both male and female, are drawn to play our "little game" (as shown in figure 1).

It is not hard to imagine the thought processes at work here, however subliminal. With no interviewer to engage them in a conversation and probe until the topic is exhausted, it is easy to see how they may answer the usual standard question. They give, for example, four answers, and then move on, feeling that they have satisfied the task ─ a classic case of 'satisficing.' But, when we present them with a game, they are challenged to give more and fuller information. Gamification drives them to overcome the satisficing behavior we standardly see.

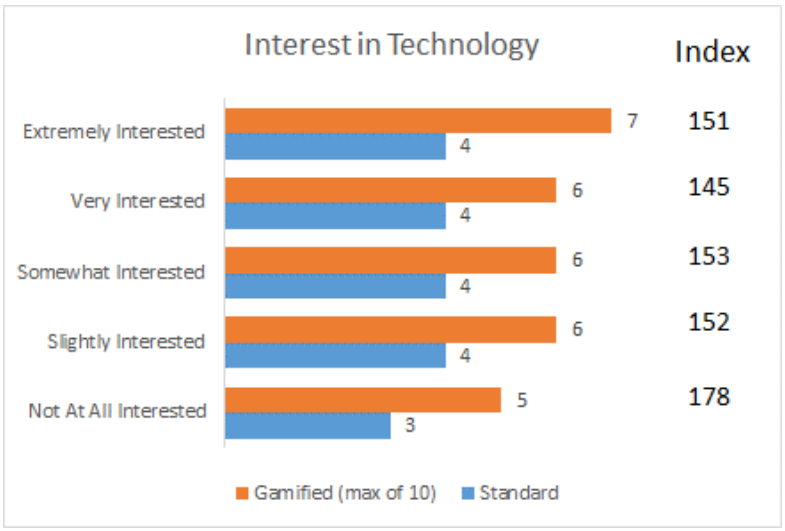
Figure 1. Average number of brands by age and gender



Base: all countries
** Index is the percentage change in the average number of brands from the standard to the gamified question = average number of brands in the gamified question/ average number of brands in standard question * 100%*

Saliency and interest in the topic is more of a driver of satisficing than age. This is clearly evidenced by the largest increase in the number of brands given by those who are least interested in the subject. It is they who give the fewest answers in the standard question (see figure 2).

Figure 2. Average number of brands by age and gender

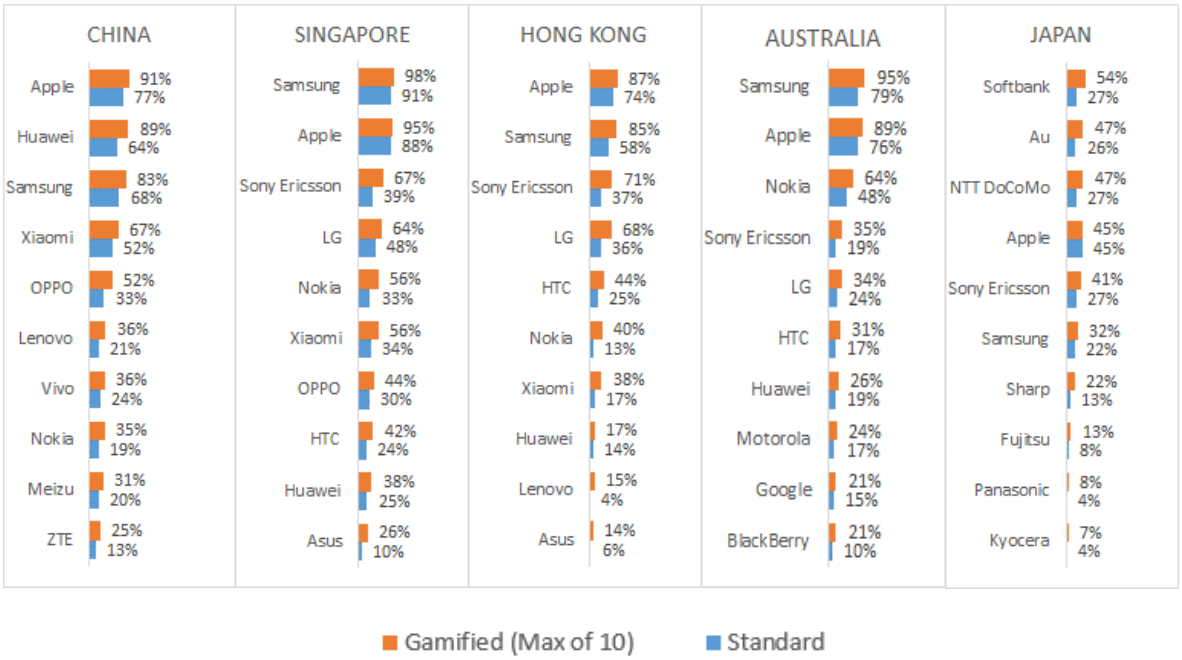


Base: all countries

These additional brands were not obscure "extra" brands. That would make no difference to the interpretation we put on the data we derive. But, rather, more mentions of the best known brands. Figure 3 shows that the level of awareness across markets substantially increased for almost all of the top brands in the gamified question. It clearly shows that people are satisficing in the standard question. When we turn the question into a game, people were roused to give an optimal response, as guided by the rules that we set.

The impact on data quality is significant. Any researcher who is trying to measure the brand's competitive performance using the standard question can run the risk of misrepresenting their client's advertising and marketing effort. This can also have a material implication for the brand. A key player like OPPO in China, for instance, might go for a different media strategy given the substantial higher awareness level for its brand as well as its key competitors.

Figure 3.. Percentage awareness – Top 10 brands



People like playing games. We are all wired to achieve the goal that we are challenged to reach, regardless of culture, age and gender. Gamification offers us an exciting opportunity to engage our participants to give us better and richer responses, by tapping into people's natural "achievement motivation."

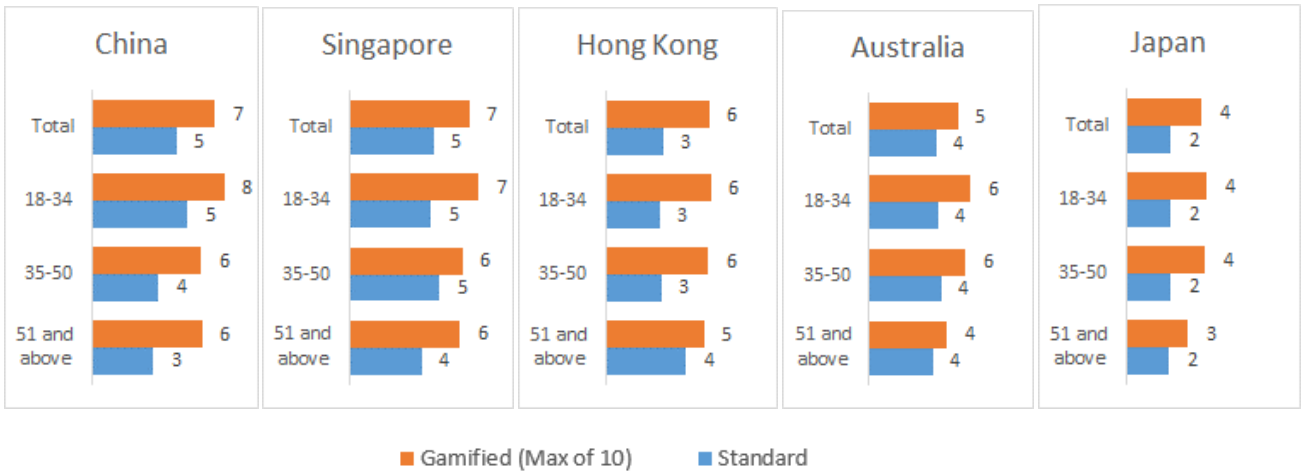
Do all cultures play games to the same extent?

While playing games can be considered a universal activity, cross-cultural variability exists. Looking back at our earlier findings (table 1), we can see that not all countries "play" to the same degree. Australia, for instance, shows a lower increase in the number of brands mentioned in the gamified question compared to Hong Kong and the rest of the markets.

Research suggests that reasons for differences in cross-cultural response styles (or the different ways people from different countries answer surveys) can be *dispositional* – related to individual characteristics such as age, gender, personality (or interests), culture, language capability or *situational,* i.e. related to characteristics such as the ambiguity of questions or language of the questionnaire (Harzing, 2006[11]). In looking at the inter-country differences, we therefore explore these same sets of factors that can also potentially influence one's willingness to play by the rules.

As we look into the demographics, there are no gender differences in any market. Both males and females play the gamified question to the same degree. Analysis by age is less clear cut. Looking at the results in figure 4, we can see cross country differences in the level of response amongst different age groups. In some countries like Singapore and China there is a large increase in the average number of brands given by the older group in the gamified question. However this is not so apparent in Hong Kong, Australia and Japan, where the percent increase for the older group is not as big as those in the younger group.

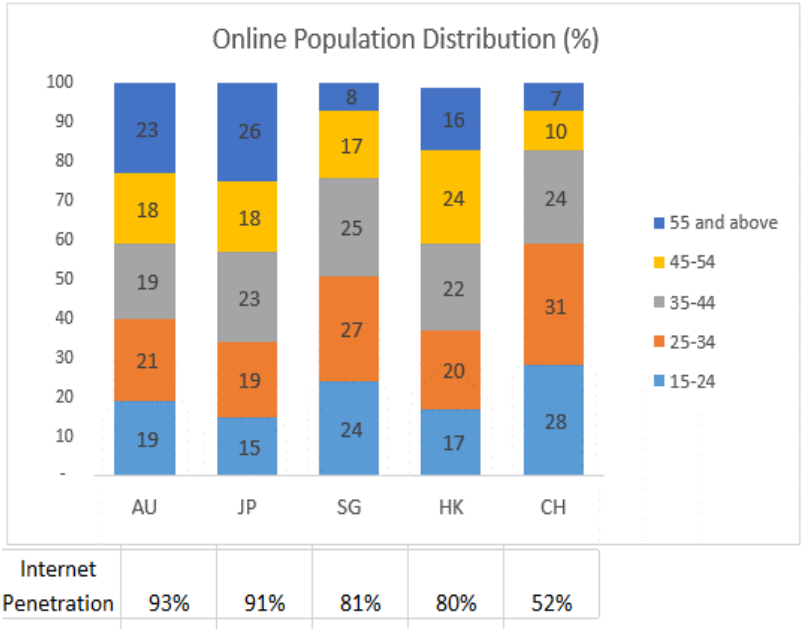
Figure 4. Average number of brands by country and age



To understand the age differences between countries, we need to keep in mind that this is a function of the internet penetration as well as the rate of technological adoption within the market. Typically, in most unsaturated markets, the online population will show a skew towards younger people. This means that older people who are online tend to be less representative of their own age cohort. This can impact the data variations that we are seeing in the data, where older people in an emerging internet market like China *seem* to play more compared to their counterparts in Australia, when in reality, it is a special type of older person who is online in China and they might be totally different from the rest of this older age group.

In addition, the vibrant tech community in China and Singapore as a whole is a major factor in the high level of response that we see across age groups. With the ubiquity of mobile phones and the increased data connectivity in these markets, it is not surprising to see that older people who are online are slightly more willing to "play" the game compared to the rest of the markets, particularly when the topic is mobile phones.

Figure 5. Online population by country and age



Source: Comscore and Internetworldstats

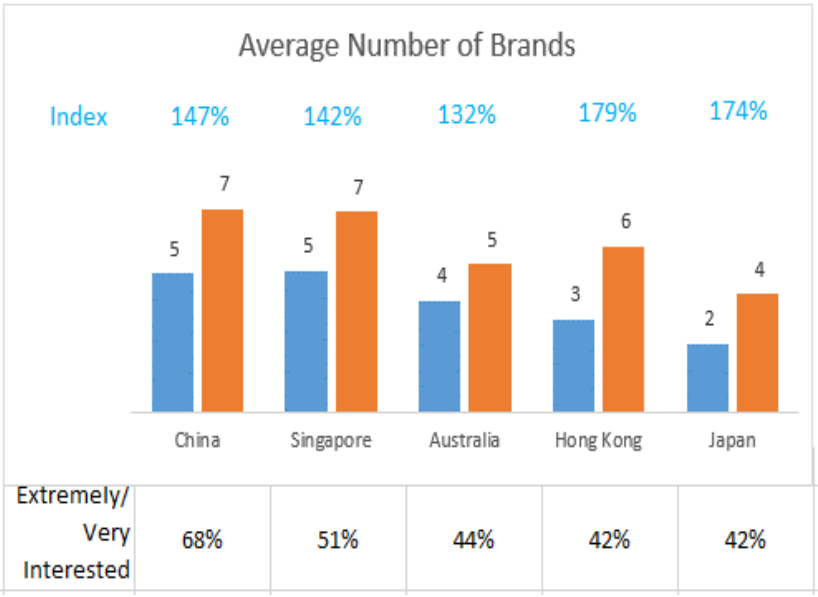
What is clear, and perhaps more telling, is the varying levels of interest in the subject across markets. As earlier mentioned the level of interest in the topic drives people's willingness to answer (or to satisfice) in the standard question. In figure 7 we can see that China and Singapore appear to be more interested in the topic, consequently showing less satisficing behavior when answering the standard question compared to Hong Kong and Japan. This then impacts the extent that they *can* play the game.

However, to show that the variance we see is mainly due to differences in tendency to satisfice, we have to assume that everyone actually has the same amount of knowledge. A cursory examination of sites of major mobile phone and service providers in the different countries (China Mobile, SingTel, Telstra, NTT Docomo) shows, however, that there are more mobile phone brands available in China and Singapore compared to Hong Kong, Australia and Japan. Therefore, we might expect the underlying sum total of knowledge to be higher for China and Singapore than in the other countries.

If we take the gamified data as the truth, that people actually know more than they say they do in the standard question, and the gamified question reveals the full sum total of knowledge, then the extent to which people apparently play the game depends on where they start from (the tendency to satisfice as related to their level of interest) and where they can go to (the ability or knowledge and willingness to play by the rules).

This then leads us to the question – just how far would they play by the rules?

Figure 6. Interest level and average number of brands by country

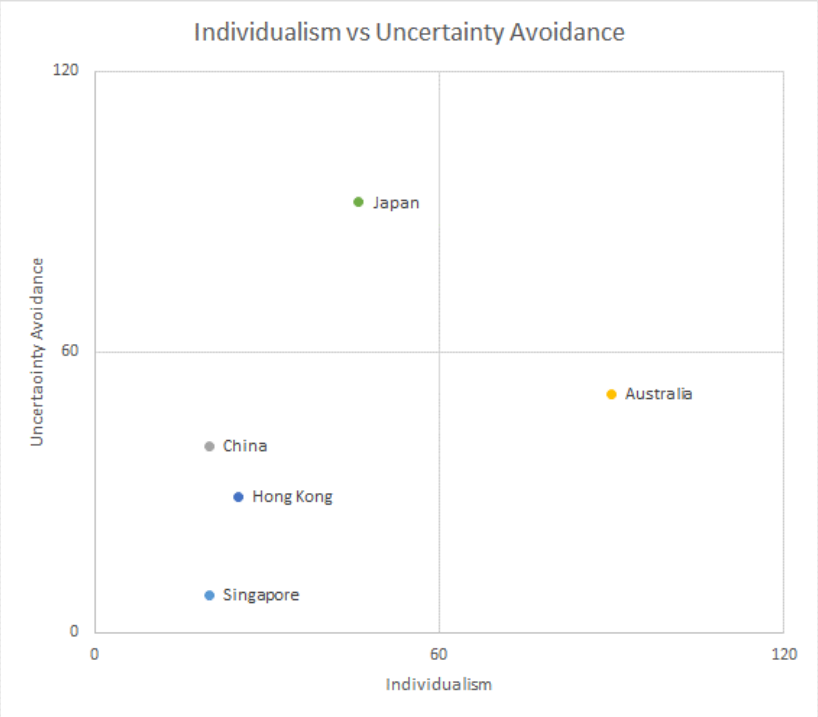


To explore the predilection of different cultures to abide by the rules, we take Hofstede's cultural dimension theory. In particular we focused on two of his dimensions - Uncertainty Avoidance (UAI) and Individualism-Collectivism (IDV) as potentially relating to rule observance and game-playing. According to Hofstede, cultures that are high on Uncertainty Avoidance look for stiff codes of behaviour or guidelines, while societies with a lower UAI score tend to impose fewer regulations. Meanwhile, Individualistic societies are more focused on "I" than Collectivist ones with their emphasis on "we".

Exploring the different cultures through this framework, we examine whether cultures that are high on Uncertainty Avoidance and highly collectivists are more compliant when rules are set. We also look whether societies that are more individualistic play harder, as they ought to be more motivated with the possibility of winning.

Looking at Hofstede's country level scores, we can see that Australia is a highly individualistic country compared to Singapore, China and Hong Kong who each share the same collectivist culture. While Japan is not as collectivist as the rest of the markets, it is certainly one of the most "uncertainty avoiding" countries, possessing a strong regard for rules. With that in mind, we dive deeper into the data and unravel very intriguing results.

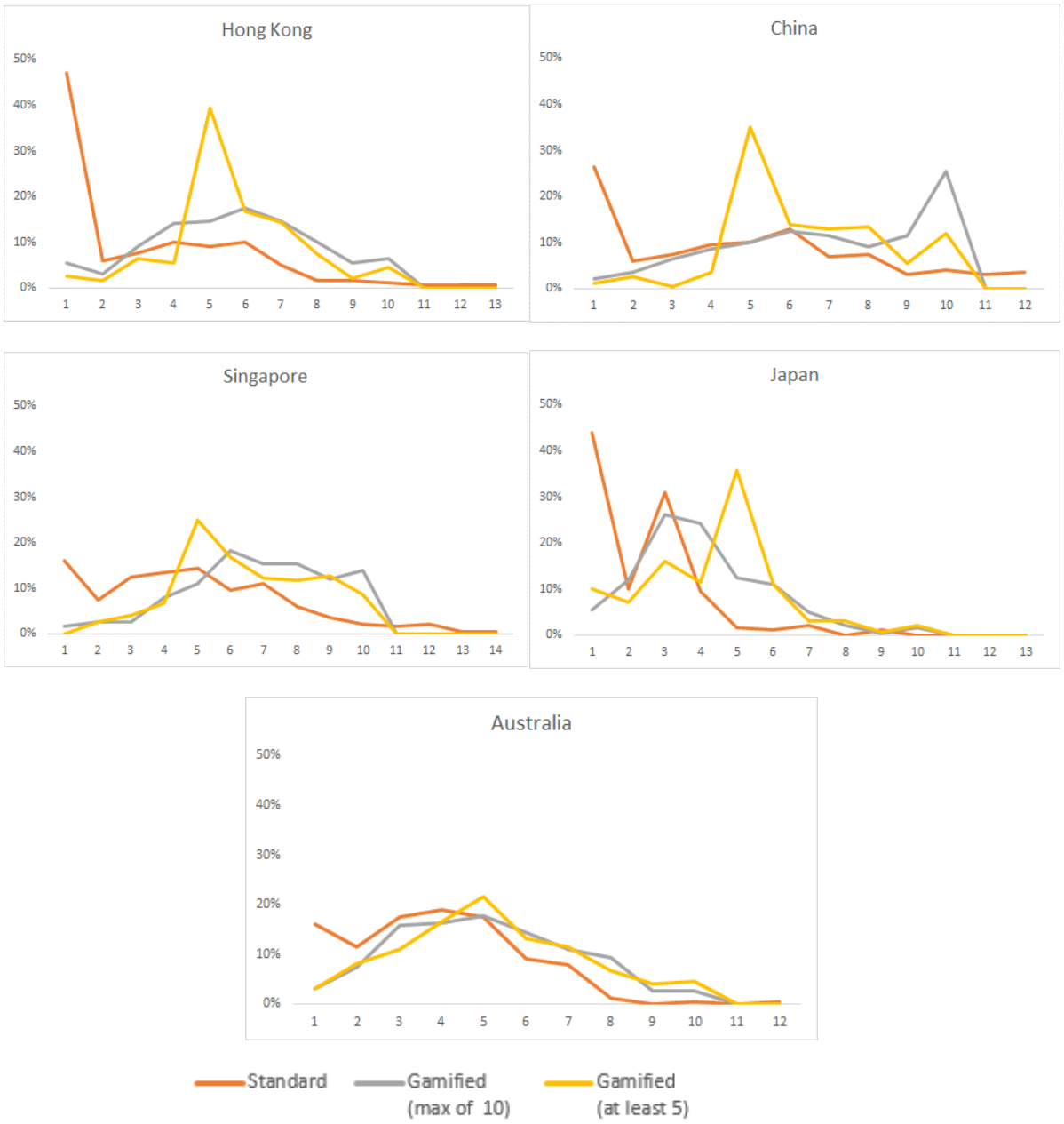
Figure 7. Individualism vs. uncertainty avoidance



Our study reveals that people in collectivist and "uncertainty avoiding" countries like Hong Kong, Singapore, China and Japan are indeed more conforming to the rules compared to an individualistic country like Australia. This is clearly evidenced by the results when we changed the rule from "a maximum of ten" to "at least five." In figure 9 we observed a noticeable spike in the number of people giving five brands only for all markets except in Australia. While there are peaks and a slight spike at five, Australia definitely shows more consistent distribution across the standard and both sets of gamified data.

This shows that the extent to which the different countries play the game relates to their culture. People in collectivist cultures tend to pay more attention and conform to the rules that we give. When we set the conditional rule: "You need to get at least five", they get to five and then they stop. Whereas people in an individualistic society tend to go on and 'win' the game, writing down as many as they can, and take less notice of conforming to the rules.

Figure 8. Distribution of number of brands by country



While gamification proves to be a powerful tool in eliciting engagement and more accurate response, it also demonstrates how the rules that we set can dramatically change the way people play the game. In the charts shown in figure 9 we can clearly see the difference in the outcome based on the rules that we used. In our first experiment, “Max of ten” is the rule that mostly gave the full set of answers, whilst “At least five” is not a competitive objective as much as a license to satisfice (as shown by the spikes in the number of people giving five brands only in the charts in figure 9). If we closely look at the standard and the gamified (max of ten) data, we can see the different peaks in the distribution and how the rule “Max of ten” closely resembles the standard data especially if we take out those people who satisfice and give just one brand in the standard question. This perhaps tells us the real distribution in each of the markets which further suggests how people in a collectivist society are more conforming when explicit rules (to a maximum of ten) are set than when it is implied as with the standard question (write down all you can think of). Meanwhile Australia, being highly individualistic, is more consistent in playing the game, showing less satisficing behavior in the standard question.

Further analysis of the data shows that those people who name ten brands actually do give real brands. However, there are cases where the participant fills in one box with a non-mobile brand, like a name of a mobile service provider, to make it to ten. This is an excusable 'cheat' and does not affect data quality. There are also some people who add in brands that are out of the category or plain gibberish just to 'win', but they are very much in the minority at 3% – 5% percent in the different markets. Again this is not a problem, these same types of people exist in the standard data set; the game has not caused this behavior. What is more unsettling is the large tendency to satisfice in the standard question. The huge difference in the number of people who gave one brand across the three treatments (standard vs max ten vs at least five) illustrates the impact that gamification can have on data quality across all markets. It is worth noting that this was not a particularly long survey, in fact it only asked this particular set of questions, but nearly half of the participants in Hong Kong or Japan just gave one brand when in reality they must know many more than that.

Table 1. Distribution of number of brands – Hong Kong and Japan

	HONG KONG		
No of brands	Standard %	Gamified (max of 10) %	Gamified (at least 5) %
1	47	6	2
2	6	3	1
3	8	9	6
4	10	14	5
5	9	15	39
6	10	18	17
7	5	15	14
8	2	10	7
9	2	6	2
10	1	7	4

	JAPAN		
No of brands	Standard %	Gamified (max of 10) %	Gamified (at least 5) %
1	44	5	10
2	10	12	7
3	31	26	16
4	10	24	12
5	2	12	36
6	1	11	11
7	2	5	3
8	0	2	3
9	1	0	1
10	0	1	2

As researchers, our inability or unwillingness to act and take a more creative approach in designing our surveys clearly presents a serious threat to data quality. Not only may the data be wrong but the inference that one can draw from it may be far from the truth. To infer that old people do not know anything about mobile phones can easily result in missed opportunities for brands and marketers or wasted resources. In fact the real inference is that old people know about the category but they are simply not that interested in the topic, certainly not interested enough to share what they really know and perhaps think.

Everybody plays games but we do it in different degrees. How well the game is played is defined by the rules that we set.

What is a good rule to set?

In the 1970s Hungarian psychologist Mihaly Csikszentmihalyi introduced the psychological concept of "flow" – a state of heightened focus and immersion with the activity at hand[12]. In his theory, Csikszentmihalyi explains that a person's skill and the difficulty of a task interact to result in different cognitive and emotional states. When challenges are too difficult and one's skill level is low, people become anxious. Alternatively, if the challenge is low but one's skill level exceeds those challenges, people become bored. According to his model, a balance must be struck between the challenge of the task and the skill of the performer to achieve a flow state. It further postulates three conditions that a task must have to increase the probability of entering a flow state:

1. Clear set of goals and rules
2. Clear and immediate feedback on performance.
3. The *perceived* challenges of the task at hand and their own *perceived* skills must be matched and high

These factors of flow are said to be interrelated. To have a perceived balance between challenges and skills, it requires that one knows what he or she has to do (clear goals) and how successful he or she is in doing it (immediate feedback).

While in the flow state, people experience:

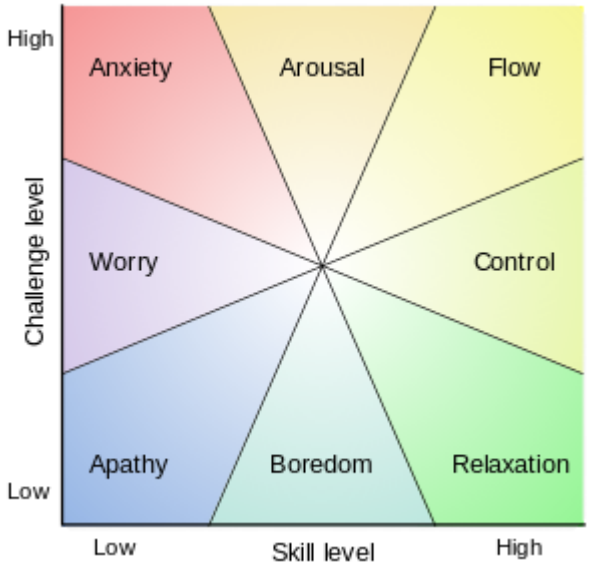
1. Intense and focused concentration on the present moment
2. Merging of action and awareness
3. A loss of reflective self-consciousness (the feeling that one is being watched or observed)
4. A sense of personal control over the situation or activity
5. A distortion of experience of time (or one's perception of time)
6. Experience of the activity as intrinsically rewarding, also referred to as autotelic experience

Flow is one of the main reasons why people play games.[13)]. It helps foster an enjoyable experience which in turn increases motivation and draw players to continue playing. Csikszentmihalyi further argues that with increased experiences of flow, people experience "growth towards complexity." As people become fully involved in using their skills to the utmost to achieve the goal, they emerge with great feelings of competence which leads them to seek increasingly greater challenges.

This concept of flow surrounds the idea of intrinsic motivation, which we often refer to in market research. It is the state in which we would ideally like our respondents to be in if we want them be fully engaged in providing their optimal response in our surveys. Given the great deal of information that we usually ask them to pay attention to, it is only sensible to understand and consider the conditions that will help them get in the flow and strive to integrate flow principles in our survey design.

To help increase the likelihood of initiating flow states, it is equally important to know the challenges to achieving flow. These include being in a state of apathy, boredom and anxiety which naturally leads to satisficing behavior. These states happen when either challenges or one's skill level or both are low (as with the case of apathy). In market research, the level of challenge and skills when taking surveys are inherently low. If we want to elicit flow experiences, we need to get them close to a state of arousal and in control. Fortunately, Csikszentmihalyi tells us how to do this: "If challenges are too low, one gets back to flow by increasing them. If challenges are too great, one can return to the flow state by learning new skills".

Figure 8. Csikszentmihalyi's flow model



This path of continual interaction between challenge and learning is also exhibited in games. In well-designed games, players perform at the edge of their competency as they are guided by clear goals and feedback[14)]. In our first experiment, it is worth noting that we also gave the respondent feedback after the question. We told them how many they wrote down (if greater than zero) and that this was a good score. This is motivational on two counts. Firstly it is feedback as to competence – we are telling the respondent that they did it well. Whether they did it or not, on an absolute basis, does not matter, they will never know and it is still motivational. Secondly it hints that someone is taking notice of what they are doing, as they do it. The fact that it is a computer does not really matter (Cape, 2016)15).

If we want the respondent to enter the flow state and have an intrinsically rewarding experience, we need to encourage the participant to increase their skill level. To do that, one must have confidence in their ability to complete the task at hand. By providing immediate feedback, we allow them to experience positive emotions which are inextricably linked with one of the basic laws of learning: Law of Effect. Murphy on his study of Why Games Work and the Science of Learning [16)] illustrates this as a law of positive feelings just as when a student who has early success with initial material will associate positive emotions with the learning experience and will be more motivated in the next steps. However, without bringing up the challenge, this is all unlikely to happen. We need to present respondents with questions that are challenging enough to maintain a competitive situation and allow them to stretch their skills, but not too difficult as to make them feel anxious, discouraged and eventually lose their interest in achieving the goal.

To further examine Czikszentmihalyi's skill-challenge balance, we look into the second part of our experiment. In this test we manipulated the number of words we asked for from a relatively easy five words to describe your favourite mobile phone brand, through a slightly less easy seven words to a frankly difficult ten words. Table 2 shows the level of compliance across markets to the different rules that we set. We can see significantly higher levels of compliance in countries that use English as their common language (Australia and Singapore) as compared to those who speak a different language (China and Japan).

The first (use exactly five words) and the second (use exactly seven words) rule appears to be challenging enough for majority of the participants to focus on and do the task, particularly in Australia. It makes them think about what it is that they really want to say about their favorite mobile phone brand using exactly five or seven words. The rich vocabulary of the English language gives them an enormous flexibility to express themselves and do the task whilst being creative and succinct in providing their response – "*It has good battery life,*" or *"Leader in Design and Technology."*

In Singapore, where most people are multi- lingual with English as their common language, the trend is almost the same. However, it appears to be more difficult as can be seen in the results from table 2. This is because of the different forms of English that range in the country from standard Singapore English, which resembles British English to a colloquial form known as "Singlish."

For China, Hong Kong and Japan it was a totally different case. In China, while we used the same rule (using exactly five or seven or ten *words*), almost all of the participants describe their favorite mobile brand in terms of concepts. Even if the question was translated into their own native language, the rule (using exactly five or seven or ten *words*) did not elicit the same response since it is not relevant in their local language. There is no concept of words in the Chinese language. All the rule has done is to make the task difficult. The majority of them are not able to comply and, most likely, end up feeling anxious for failing the task. This shows how important it is not only to communicate in a local language but also in the way the language is used in a certain market.

In Japan, translations resulted to further changes in the rules used for the country. After discussions between professional translators and employing back translations with the local team on ground, we presented the survey in Japanese, asking them to describe their favorite mobile brand with five or seven or ten adjectives. All the responses were then back translated and coded for key descriptors. Results in table 2 showed that finding five adjectives is a hard task, perhaps more difficult than using exactly five words or

characters, inducing the same feeling of "anxiety" on the part of participants. Further research should be conducted to fully explore the cultural influences in relation to rule setting.

Table 2. Distribution of compliance – Words/concepts* by country

Country	Using 5 words (%)	Using 7 words (%)	Using 10 words (%)
Australia	54	54	28
Singapore	46	43	23
China*	27	12	3
Japan*	22	14	4

In setting a rule, not only can there be differences between countries but there can also be variations within a market. In Hong Kong, for example, where the majority of the population is of Chinese descent, over 40% are bilingual and use English as their second language. All the surveys were presented to the participants in Chinese and the rule stated "using exactly five, exactly seven or exactly ten words." Participants have a free choice as to how they write in an open question, in English or in Chinese. Table 3 shows an analysis of counting either characters or words given. Amongst those who respond in Chinese, the majority describe their favorite mobile brand in characters, albeit at ever decreasing rates as the task gets harder. Where they wrote in English, compliance was lower than in Australia and Singapore. The confusion in the task instructions made the whole thing simply too difficult.

Table 3. Distribution of words/characters – Hong Kong

Hong Kong	Chinese Characters (%)	English Words (%)
5 words/ characters	71	29
7 words/ characters	68	12
10 words/characters	54	10

Given the rule was set by an English speaker, it is not surprising to see how a similar rule can create a positive experience in Australia and Singapore and at the same time pose a confusing and very difficult task in other markets. Confusion arises from their use of pictograms rather than Latin characters. Is it five words? Or five pictograms? Or five concepts? The rule simply does not fit to the structure or direction of the game in their context, as it would to an English person interpreting the game. This reveals the cultural bias and constraint that comes with the person setting the rule. Setting a universal rule is not an easy task. Thus, it is essential to seek advice from local people in devising the right rule in the market that would work best in getting better quality data.

Researchers tend to believe that people from different cultures will understand the same question in the same way. However our study evidently shows that the language and the different form it can take in one's culture can influence the way people react to the set of rules. This illustrates how crucial it is to devise a rule that is culturally relevant and will add clarity to the task and category. We need to think about what we are asking them to do in the context of the study and think through the rule that would get them to think and "play" the game.

While we understand how setting the rules concentrates the mind, we need to guard against introducing bias when defining them. We need to keep in mind that we are not playing games for game sake in market research. What we are interested in is designing a mechanic that will help the participant overcome the tendency to satisfice and give more and better quality data. Having a clear defined goal is paramount in setting the right rules that will make people focus on the task and want to win the game that we want them to play.

Summary

People are wired to want to win - to achieve the goal that we are challenged to do, regardless of our culture, age or gender. By making the survey game-like we allow participants to win at research.

While everybody plays games, we do so in different degrees. How far people will be willing to "play" depends mainly on their culture, ability or skills (knowledge of the topic and language) and most importantly the rules that we set.

It is easy to believe that people of different cultures will understand and abide by the rule in the same way. Our study clearly shows that the extent to which people play by the rules is variable depending where they start from (the tendency to satisfice driven by their interest in the topic) and where they can go to (the knowledge or skills and willingness to play by the rules).

Our study reveals the different predilection of cultures to abide by the rules as referenced by Hofstede's cultural dimension theory. People in collectivist and uncertainty avoiding countries like Hong Kong, Singapore, China and Japan are indeed more conforming to the rules compared to people in an individualistic country like Australia, who tend to want to 'win' the game.

Results from our study also demonstrate how the rules that we set can dramatically change the way people play the game and impact the distribution of responses. Understanding the society's culture and language will help us set the right rules to get people to focus on the task and want to win the game that we want them to play.

References

1. http://www.international-television.org/tv_market_data/global-tv-game-show-ranking.html
2. Predictably Irrational - basic human motivations: Dan Ariely at TEDxMidwest http://www.google.com.sg/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0ahUKEwiarNq2l63RAhXMMY8KHRr6DTAQFggfMAE&url=http%3A%2F%2Fwww.allreadable.com%2Fce2b58e&usg=AFQjCNHsNdA5ghSemgGTETm7bit99fks-w&sig2=xWab1NOgI5vqL1Gj1f45ZQ
3. Who Wants to Be a Millionaire". Gameshow Hall of Fame. GSN. January 21, 2007.
4. Zichermann G. (2014). Engaging Millenials with Gamification. TEDx . Retreived from https://www.youtube.com/watch?v=gZga0AmCGdE
5. Robertson I. (2012). Perfect Cities? The Winner Effect. TEDx Dublin . Retrieved from https://www.youtube.com/watch?v=Bdn0qcrTv0c
6. Puleston J. , Sleep, D. (2008). Measuring the value of Respondents Engagement. Innovative Techniques to Improve Panel Quality. Amsterdam, The Netherlands: ESOMAR
7. Shore N. (2011). Millennials are Playing With You. [Web log post]. Retrieved from https://hbr.org/2011/12/millennials-are-playing-with-y
8. Deterding, Dixon, Khaled, & Nacke, 2011. Gamification: Toward a Definition. http://gamification-research.org/wp-content/uploads/2011/04/02-Deterding-Khaled-Nacke-Dixon.pdf
9. 9 Keusch F., Zhang C. (2015). A Review of Issues in Gamified Surveys. Social Science Computer Review. (p4). Retreived from http://journals.sagepub.com/doi/full/10.1177/0894439315608451
10. Puleston & Sleep, 2008. Op Cit.
11. Harzing A. (2006). Response Styles in Cross National Survey Research: A 26-Country Study. International Journal of Crosscultural Management, vol 6/2
12. Nakamura, J.; Csikszentmihályi, M. (20 December 2001). "Flow Theory and Research". In C. R. Snyder Erik Wright, and Shane J. Lopez. Handbook of Positive Psychology. Oxford University Press. pp. 195–206. ISBN 978-0-19-803094-2.
13. Murphy, Curtiss (2011*).* "Why Games Work and the Science of Learning". Retrieved from http://www.goodgamesbydesign.com/Files/WhyGamesWork_TheScienceOfLearning_CMurphy_2011.pdf
14. Rutledge P. The Positive Side of Video Games: Part III. [Web log post]. Retrieved from http://www.pamelarutledge.com/2012/08/29/the-positive-side-of-video-games-part-iii/
15. Cape P. (2016, May). Gamifying Spontaneous Awareness. Research News. (p12-13).
16. Murphy, Curtiss (2011*).* Op. Cit.

The Authors

Jennifer Serrano is Knowledge Manager APAC, SSI, Singapore

Pete Cape is Global Knowledge Director, SSI, United Kingdom

The Segmentation Revival

Developing a global and universally applicable segmentation method

Alexander Wheatley • Jonathan Puleston • Jason Brownlee

Introduction: In the world of big data, why conduct segmentation research?

With the ever increasing volumes of "big data" being generated that companies either own or can access, it begs the questions why would you need to conduct any third party research to segment your audience? Why in fact are we witnessing a growth in the volume of segmentation research being conducted? In the last two years the number of segmentation studies we have fielded, we estimate, has nearly doubled. Segmentation seems to be going through a renaissance. There appear to be three underlying reasons at play.

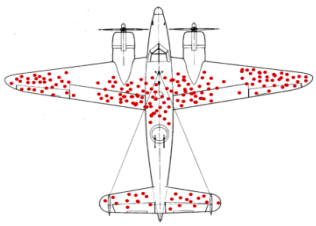
1. *Alleviating survivorship bias*:* the problem with working with your own data, no matter how much you have, is that it can give a myopic viewpoint on the wider market place. You are in effect only looking at the customers who have "survived" your sales and marketing process. We are seeing a common need for this data to be calibrated and contextualised, to understand what makes their customers different from their non-customers. This is something that a traditional piece of segmentation research can help to do.

2. *Expanding out narrow cast data*: The second problem with first party data is it can often be very narrow cast in nature, providing the what, but not necessarily explaining the how and why. We are seeing an increasing number of clients using segmentation research to help fill out a more complete picture on their first party data.

3. *Bringing data to life:* The solution that a supplementary segmentation provides is in uniting conversations and making it easier to communicate a story. Primary data can often be very dry and unengaging. If one has understood the why, and from the data has been able to compartmentalise and vividly classify, then no matter who the conversation is with one can define a market and their needs in a relatable way. From the CEO of a company to a rep of the group there is a unified understanding of who the customer is and what they want. This is only possible if one utilises a segmentation study that can be combined with first party data in a consolidated strategy.

survivorship bias:

During World War II, Researchers studying the damage done to aircraft that had returned from missions was flawed, they realised, because it only considered the aircraft that had survived their missions .



How to create a segmentation model

There is no magic formula or standard survey you can use to create a piece of segmentation research, but there is a series of best practice steps you can adopt to design an effective segmentation study.

The basic principals

The aim of an effective segmentation is to create a comprehensive map of the types of consumers in your market by identifying differences in their knowledge, behaviour, attitudes, and motivations.

The starting point

There are some common questions that nearly all segmentations need answers to:

1. *Price anchors*: Determining how much money different people are prepared to spend in different circumstances in a category
2. *Time & effort*: Determining how much time and effort people are prepared to put into their purchasing decision
3. *Knowledge*: Understanding levels of prior knowledge about decision making within category as this will have a big effect on how people go about making decisions.
4. *Established behaviours*: Where this information is available it can be really useful for understanding default and established behaviour - essentially what people normally buy or do
5. *Conscious decision making processes*: Understanding what information people use or think they will use to make purchase decisions and the metrics they use to make their choices
6. *What are the emotional & less conscious drivers of choice*: Finding out what are the less conscious drivers of choice
7. *The purchasing circumstances*: Understanding how purchasing circumstances vary and how purchase environments might influence personal choice

Mapping out the conscious and unconscious decision making drivers

The two biggest challenges are to map out the **conscious and unconscious decision** making drivers and convert them into a set of usable questions.

Most businesses will probably have a good understanding of the primary decision making drivers, particularly in the market they operate, but when designing an international segmentation we strongly recommend conducting a preliminary qual phase or quali-quant study at a cross section of markets to fully understand the breadth and depth of active factors across different cultures to ensure nothing has been overlooked.

Stage 1: A preliminary qual-quant study

A technique we recommend is a very simple short online survey with a few open ended questions like the examples below. Sent to a hundred or so people representing a cross section of each market, this can go a long way to help you understand the primary conscious drivers of purchase.

The types of questions to ask…

1. What are their general attitudes towards the category
2. What do they think about when making a purchase decision
3. How do they normally go about making their decision
4. Examine a recent purchase to find out the reasons they chose what they purchased
5. Ask questions about what stops them from buying a certain product
6. Find out what helps them make decisions

The advantage of doing this quantitatively rather than qualitatively is that you get a sense of the relative significance of factors and you avoid feeding too many phantom factors that might be peculiar to single individuals - if a factor has not been mentioned by more than one person in 100 people, it's unlikely to be that significant. The breadth of questions should still ensure nothing is over looked. This process should give you a broad list of purchase drivers and things to think about, and by asking this quantitatively you will get a sense of how significant each factor is by how often it's mentioned.

Table 1 provides an illustration of the type of insight we got back from a conducting a small quali-quant study exploring people's attitudes towards cooking in the UK & India.

Table 1.

	Mentions
Expressing a passion about cooking	40
Expressing enjoyment about cooking	120
Cooking being an opportunity to be creative	20
Discovery - chance to experiment and try new things	18
A form of relaxation	17
Health factors	31
Family mentions	25
An opportunity to nurture themselves and their family	8
A means of expressing their feelings of love towards their family	6
Time related/ the time it takes & time they have to do it	60
The work involved	40
Difficulty in deciding what to cook	8
A cause of annoyance over the task involved and mess it creates	8
Conflicting feelings - love hate relationship	15
A cause of stress not being able to do it well	3
How it stimulating the senses of taste and smell	8
Neutral thoughts about cooking	16
Negative feelings / it's a chore	25

There are some factors that the majority of people will mention, in this case talking about their enjoyment and what they got out of the experience, for which you may need to build up a basket of questions, while others are more niche factors which may be smaller but might well be important choice differentiators for marketers to focus on.

It is vitally important that you conduct this preliminary research within a cross section of different countries or you risk not picking up on some significant factors. The number of negative feelings mentioned about cooking in India were minimal, for example, and in the UK nobody mentioned anything directly related to nurturance yet which was a more significant factor in India.

Stage 2: Understanding the 'less conscious' drivers of choice

Conducting a qual pilot will provide you with a good list of conscious decision factors but you also have to consider the less conscious, emotional drivers alongside these. You cannot rely on research audiences being able to self-identify these. So to pull them out, we recommend going through a cheat sheet list like the one in table 2, which maps out what academics and researchers from various sources have identified as the most significant unconscious factors driving choice. Think how they apply to and how relevant they are to the category under consideration. We by no means claim this to be an original or comprehensive list, and any marketer will probably have a range of their own unique factors that would need to be added to this list, but it's a good starting point. The following are examples of emotional choice drivers:

Security	How important is it to make a safe choice, willingness to take a risk?
Mastery	Will the product purchased make me better at performing a task? How strong is my desire for mastery?
Empowerment	Will the product empower me, make me more capable, am I looking to be empowered?
Achievement	Will my choice give me a sense of achievement, make me feel proud?
Esteem	Will my choice raise my standing, my level of social esteem, how important is this?
Belonging	Will it help me connect with others, how strong is my desire to be connected with others?
Conformity	Will this choice help me conform or the opposite – will it make me stand out from conformity
Nurturance	Will buying this make me or others feel good
Health & wellbeing	Will this product satisfy my desire for myself and my family to lead a healthy life?
Engagement	Will this product absorb me, make me feel more productive?
Loss aversion	Am I driven by the need to hold on to what I have got?
Excitement	This product will excite me?
Aesthetics	Will this product deliver aesthetic pleasures? How important is this?
Urgency	How much am I driven by a sense of urgency to make a decision?
Desire for change	How eager am I to try new things? Do I get bored doing the same thing?
Mood	Is my purchase driven by the mood I am in?

To help you understand how relevant each factor is, it's useful to think about what differentiates the category from others. Here is a list of common category differentiators that can help guide you through this thinking process. Again, this list is by no means comprehensive but should act as a useful starting point.

Examples of category differentiators:

- *Considerations to do with the purchase cycles***:** Long purchase cycle consumers will have less prior knowledge and experience compared to products we buy regularly
- *Is it a fixed/routine purchase (like insurance) or an impulse purchase (chocolate):* The impulse triggers are going to be far more important for impulse purchases
- *Level of innovation within the purchase cycle:* How much has changed between purchase cycle (washing powder low, mobile phones & music high). With high cycle innovation understanding the changing needs is going to be far more important.
- *How established are the choices*: (washing powder 100 years, 3D printers 0). With unestablished choices, awareness and price expectations are less set, there are no established patterns of behaviour.
- *How big is the pool of choice* (washing powder 5, wine 500) – with large choice sets it's important to map out how people filter through the choices
- *How much does the purchase project someone's identity & values* – e.g. clothes high, toothpaste low
- *Access to products* - do they have to be searched out, can they be easily browsed and compared (confectionary - compare, car - harder)
- *What senses are involved* (FMCG - eyes, music - ears, car - eyes, electricity supplier - mind)
- *How strong is the role of the analytical mind vs. emotions***:** When choosing a car there is a lot more to analytically consider compared to choosing a chocolate bar.
- *Is it a personal or collective choice* (toothpaste - all , shoes - me)
- *How complex is the choice* (finance - high, FMCG - low)

What this process will deliver back is a long list of factors you can turn into questions. You will have to make some important choices about what factors to track in the full segmentation, but don't worry too much at this stage how long this list is – just try to make it as comprehensive as you can.

Stage 3: Thinking about how to ask the questions

Having identified all factors that you think will be important for understanding the behaviour of consumers, how do you now convert these into survey questions?

To effectively compare the relationship between one factor and another, segmentation experts really prefer scale based measures. There is a great temptation to go about setting up your segmentation survey by creating a great big long list of Likert scale attitudinal statements that you ask people to agree or disagree with.

This is probably the single biggest mistake you could make when designing a segmentation study, especially if you are constructing it to run across different markets.

Attitudinal questions require conjecture that is subject to all sorts of cognitive biases that vary significantly from market to market. You will get completely different answers simply based on the way people tend to answer such questions in each country. This issue has been explored in several previous ESOMAR papers, e.g. *Dimensions of Data Quality,* Puleston & Eggers, ESOMAR 2012.

About the primary research we conducted for this paper

To fulfil the purposes of this paper (to illustrate different ways of asking segmentation questions), we conducted a small segmentation research experiment of our own in two different countries, the UK & India. These two countries were selected because they represent the extremes in how respondents answer survey questions in different cultures.

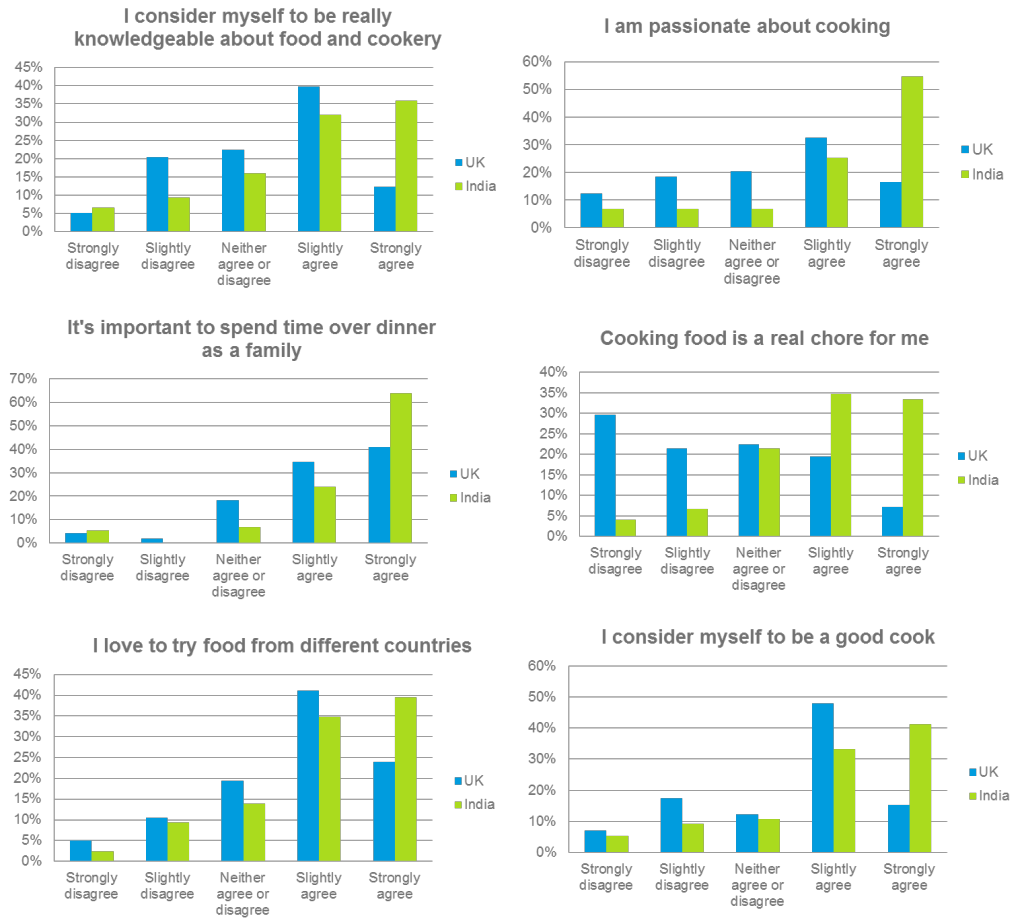
The topic we chose for this experiment was food and cooking. We polled 600 people in each country and asked a range of typical segmentation question about this topic. We developed the questions using the methods outlined in this paper, by conducted a small preliminary quali-quant study to identify the key differentiating attitudes. We asked the same questions in different ways to split cells of respondents in each country to compare responses.

The problem with agreement scales

Presented visually in figure 3 is the issue in a nutshell about asking basic attitudinal questions - the answers you get from the UK and India are so different it is almost impossible to compare them.

Acquiescence bias is strong in India, and less than 10% of answers are in the negative side of the scale. In the UK the balance of data is also often highly skewed, to the mid-top and bottom end of the scale, making it difficult to pull apart different distinct segments.

Figure 3.

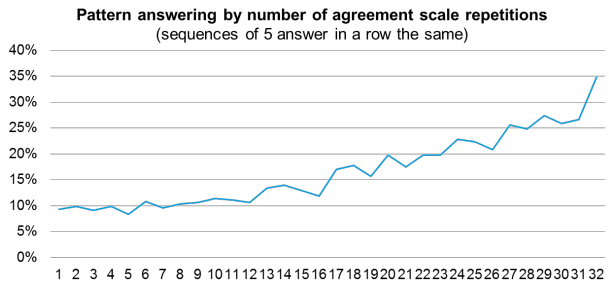


Another important factor to consider is the sheer boredom respondents experience when answering repetitive banks of agreement statements like this – their attention very quickly wains and so too does the reliability of the data. Again this issue has been highlighted in several previous ESOMAR papers. Figure 4 is a chart taken from one of these papers (Puleston & Sleep 2011) showing that after the 15th repetition, straight-lining increases rapidly.

So how do you tackle these issues?

There are many alternative ways to ask questions like these that can deliver back more cross comparable data.

Figure 4.

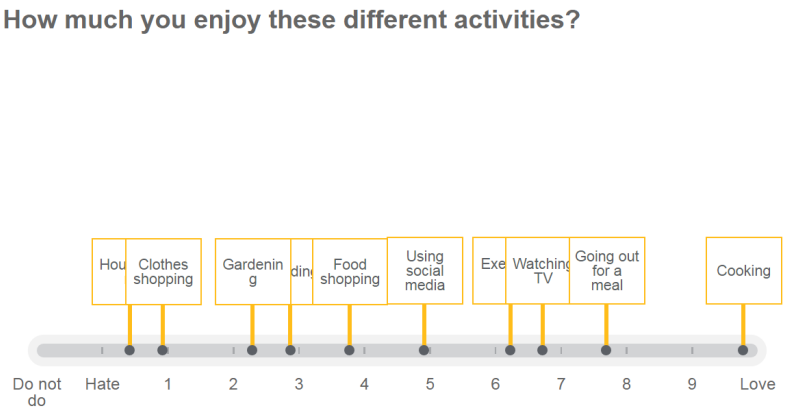


Firstly dealing with the different levels of enthusiasm
Respondents in India, as you can see, are far more enthusiastic in their answers to nearly every question than the British, and particularly to the question about how passionate they are about cooking, a statement that over 50% agreed with.

One of the best ways to deal with this problem is a simple calibration process at the start of the survey to benchmark and differentiate respondents' levels of enthusiasm.

In our experiment we use this approach to measure relatively how much people like cooking compared to a range of other activities on a rating scale using a popular question format we have devised called the flag drag and drop (see figure 5).

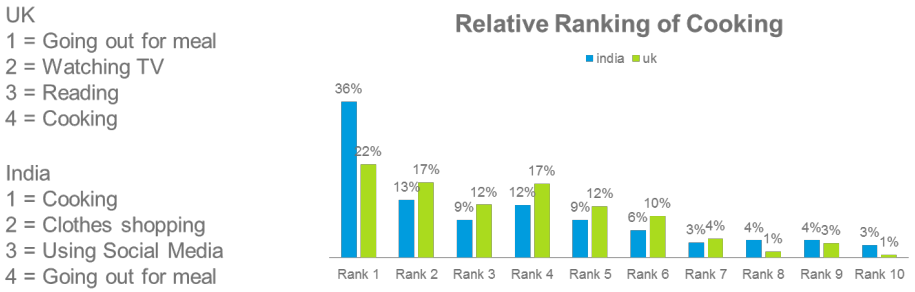
Figure 5.



The data can be used to determine just how much people really do like cooking compared to other activities, but also to help generally calibrate all the respondents other answers.

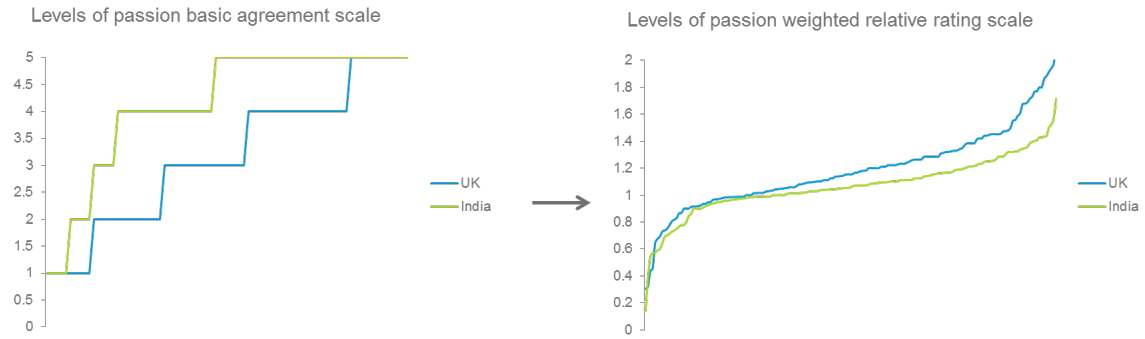
When we analysed the data in this way, as it turns out people in India do seem to be a lot more passionate about cooking compared to people in the UK when you compare their relative enthusiasm with other activities, nearly 36% ranking cooking overall as the activity they enjoyed most compared to only 22% of British people. (See figure 6.)

Figure 6.



For the segmenter, this approach provides a far more neaunced, evenly distibuted bell curve of responses too - so relative levels of passion about cooking can be more acurately compared. Figure 7 is a comparison of the spead of data when passion levels are asked on a 5 point agreement scale vs. an individually weighted comparative rating approach. What you see is a smoother differentiation.

Figure 7.



The issue with measuring something in this way is in effect this adds a whole lot of extra questions to the survey that don't tell us anything specific about cooking, so it's not practical to use this approach for every question in the survey. So, how can we ask questions in more economical ways that will reduce the levels of acquiescence bias?

A shift to double ended bespoke scales

Our first suggestion is to change from agreement scales to using more bespoke double-ended scales where neither end of the scale is specifically positive or negative. This means respondents are not being asked to 'agree' with things.

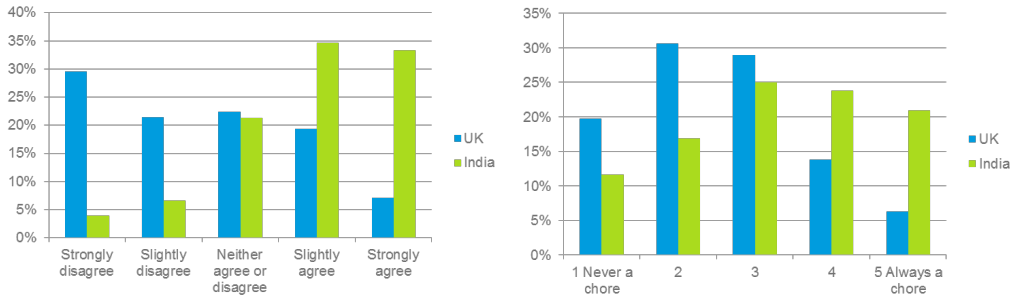
Here is an example, compared to a normal approach: *Cooking is a chore… Strongly disagree - Strongly agree*. Where as you can see from figure there are significant cultural differences between the UK and India in the answers to this question. Below is an alternative way of asking this question using a more bespoke double end scale with less inherent judgement.

What is your attitude towards cooking?

I never find cooking a chore - I always find cooking a chore

This change by no means eradicates the differences but it certainly harmonises the answers more closely, reducing the level of top box agreement in India and high levels of negative agreement in the UK (figure 8).

Figure 8.



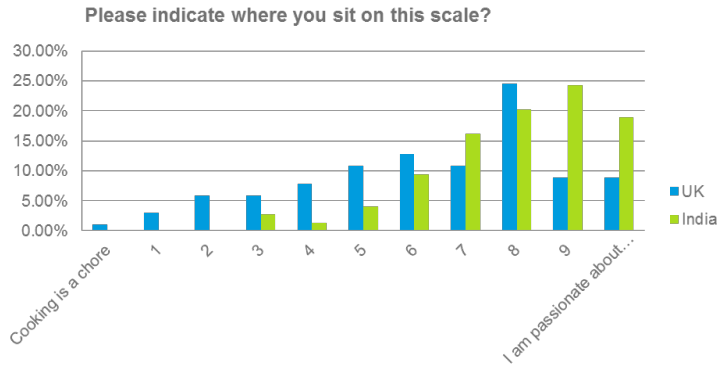
Using conflicting end scales

Another potential solution is to place two conflicting choices at the end of the scales. For example, a lot of people say they find cooking a chore but a similar amount say they are passionate about cooking - which emotion is dominant?

I find cooking a chore – I am passionate about cooking

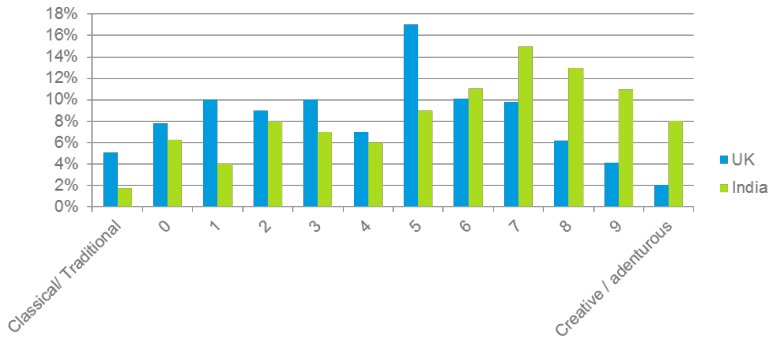
The problem with this approach is that the dominant feeling can create an even more lopsided response, in this case passion wins over chore (figure 9).

Figure 9.



It's not to say that conflicting scales cannot be extremely useful for other types of differentiation. Take the example shown in figure 10 where we use this approach to explore the type of cook people felt they were, and where we found a more interesting spread of opinions.

Figure 10.



Our advice would be to use this approach with some thought and not without some piloting (an important topic we will come on to later).

Using custom anchor points on range scales

A more effective approach is to create a bespoke scale with custom anchor points that reflect a fuller spectrum of feelings people may have. The technique is adapted from kids' research where young people have a tendency to only use the extreme end of scales when answering questions, known as the "wow-yuck" problem. Where all the answers are clustered at one end of the scale, a solution is to stretch the scale out and present a clear range of customised scale anchor points that define the different points.

So instead of asking a basic scale, "*I am passionate about cooking" Strongly disagree – strongly agree*

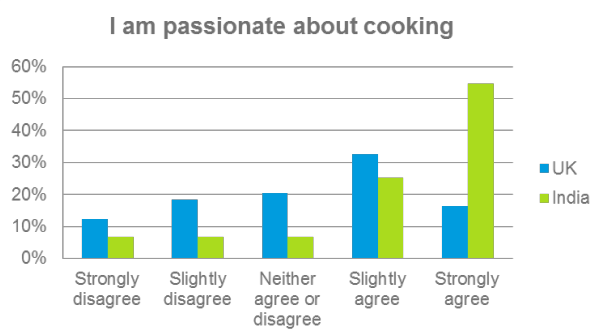
Change to →

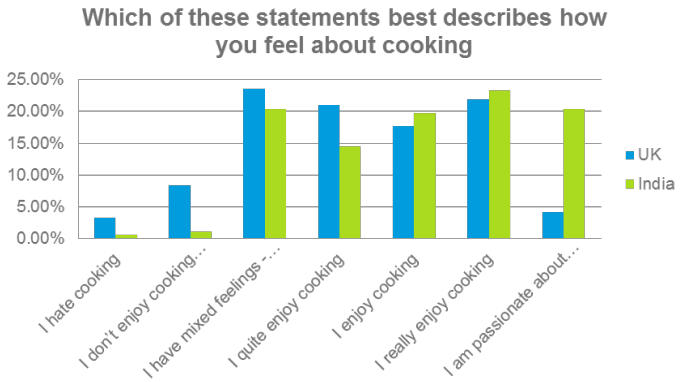
Which of these statement best describes how you feel about cooking?

1. I hate cooking
2. I don't enjoy cooking very much
3. I have mixed feelings - sometimes I enjoy it sometimes it's a chore
4. I quite enjoy cooking
5. I enjoy cooking
6. I really enjoy cooking
7. I am passionate about cooking

This change as you can see give more even and consistent distribution of answers (figure 11).

Figure 11.



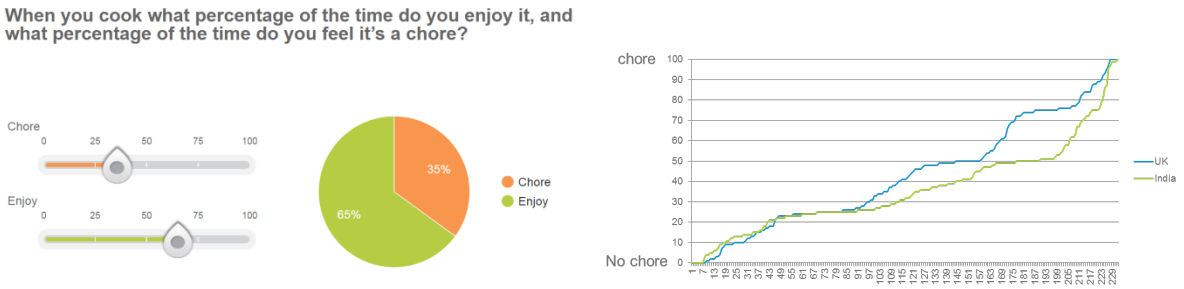


Changing the whole style of the question

One of the reasons y many people have difficulty answering a question like, “Do you find cooking a chore” or “I am passionate about cooking” and similar questions to this, is because there is often no definitive answer, they might enjoy cooking some of the time and other times it’s a chore so the real answer is “it depends!”

In these circumstances one alternative approach is to completely change the question to an allocation style process to quantify how much of the time people find cooking a chore and much time they enjoy it (figure 12).

Figure 12.



Whilst in a sense it is an abstract question and might be difficult to allocate a precise figure to, this delivers back a wonderful spread of data that the segmenter can use to evenly differentiate the sample.

Using comparative scales

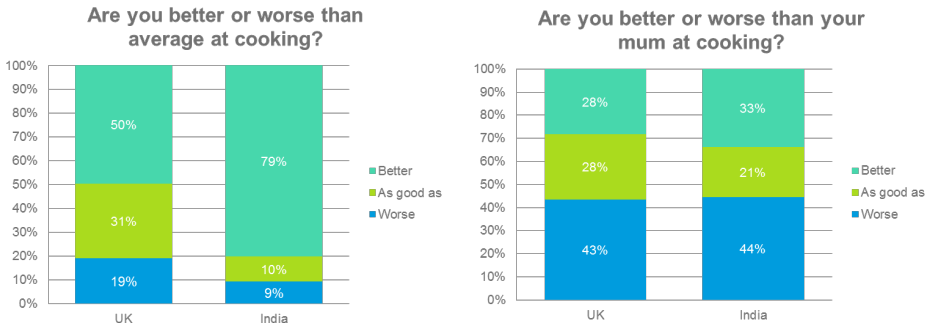
When asked “Are you a good cook?” the vast majority think we are as good as or better than other people, which is not very useful information for a segmenter.

A technique to gain a more balanced set of answers is to ask people to compare their cookery skills to someone they know like their best friend or mother. You first ask them to assess their skill and then ask them if they think they are better or worse (figure 13).

Figure 13.

As you can see the answers become a lot more consistent (figure 14).

Figure 14.

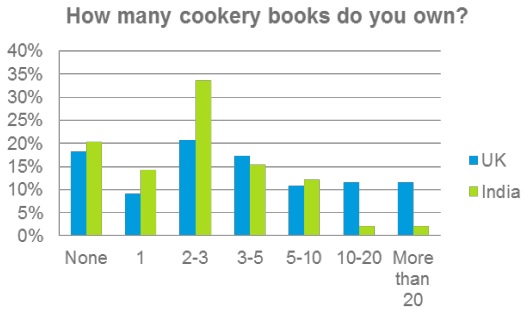


Using behavioural scales

Another technique is a shift to using behavioural scales – measuring a behaviour that is a manifestation of the attitude. (See figure 15.)

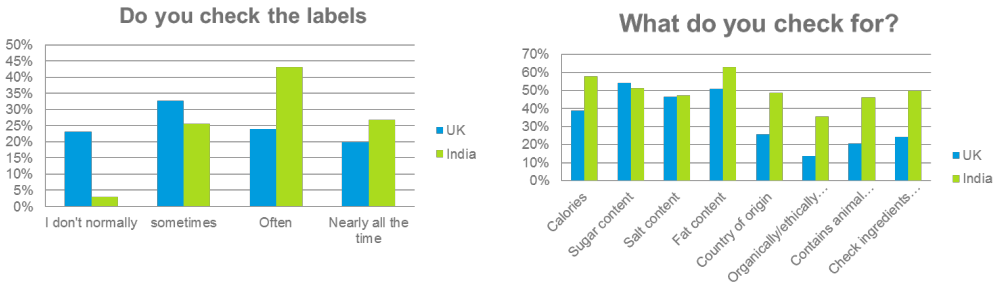
I say I am knowledgeable about food and cooking but how does that manifest itself, have I got lots of recipe books, do I watch lots of cookery programs, how much do I cook. By measuring these things you can get a potentially more truthful answers that can be more reliably crossed compared.

Figure 15.



I say I try to eat healthily but how does that manifest itself? When they go shopping do they look at the nutritional information on the label, what details do they say they look out for? Would they pick a healthy meal option from a presented list of choices? (See figure 16.)

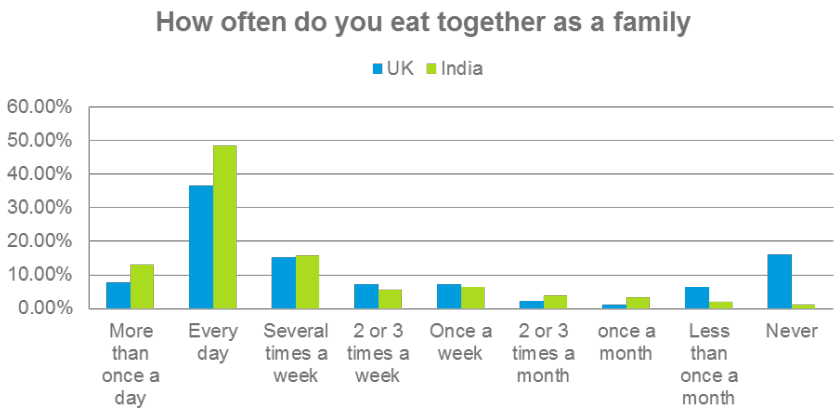
Figure 16.



I say that I think it is important to spend time over dinner with my family. As you can see from the earlier chart, nearly everyone agrees with this statement in both the UK and India as there are only 5% who disagree.. This question is therefore of little value to a segmenter, as it does not help to differentiate the population. A behavioural way to measure this might be by asking how often people actually eat together with their family.

When you ask how often you eat together as a family, much clearer segments emerge. Figure 16 illustrates that a significant group of people in the UK, for example, t do not tend to eat together.

Figure 16.

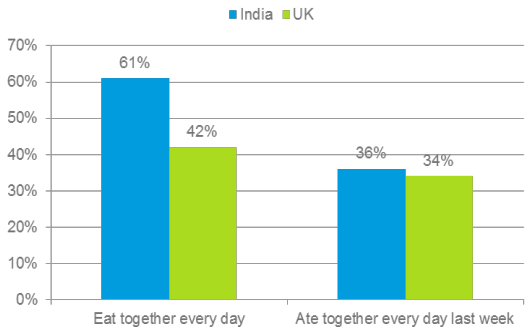


Try to ask about recent behaviour not general behaviour

However this does not completely solve the issue. Social desirability biases will also impact on behavioural questions. When you are asked how often you do things you are prone to over or under estimate based on the social desirability of the activity. Declaring how much you drink every week is a classic example of something we under estimate. It is better where possible to ask more concrete, specific questions about recent activity than about general behaviour.

So instead of asking generally how frequently you eat with your family it might be better for example to ask how many times they have eaten together over the past week. Figure 17 illustrates the dramatic difference and how similar the answers can become between the two countries when asked in this way.

Figure 17.



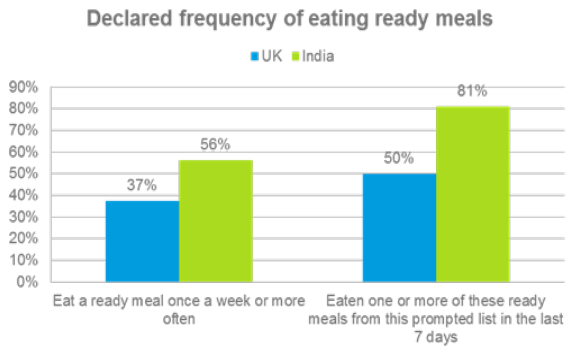
Adding memory prompts to behavioural questions and breaking into smaller thinking chunks

It's not just a social desirability bias that can corrupt the answer to a behavioural question though, it can often be difficult to remember and calculate how often we do things.

If I were to ask you how many pre-prepared/readymade meals you have eaten in the last week, many of us simply cannot remember. There are also lots of different things that could be classified as a pre-prepared meal, so I might have a preconceived idea of what a pre-prepared meal is, does it include frozen pizza and what about soup? It is even harder for me to then accurately assess how regularly I eat readymade meals generally (figure 18).

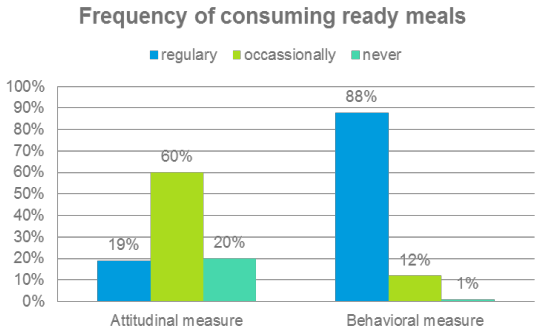
Asking a simple preliminary question about what types of pre-prepared/ready meal you have eaten recently will help remind people of the different types of readymade meals and will then make it easier to remember how many readymade meals they have consumed overall. The result is an increase in the numbers.

Figure 18.



If you were to go one step further you can measure how regularly people consume each of these different types of meals and hence add up overall how frequently they consume one or more types of meal. This can then be compared to how much they say they consume using a basic attitudinal range question. Doing this you begin to see a big gulf emerge – a leap from 20% to 90% of people who regularly consume at least one type of ready meal (figure 19).

Figure 19.



Killing 2 birds with one stone

Although you are increasing the length of the survey by adding these extra questions, in this case we are also learning something else from asking this. We learn about the character of food choices each person makes, which is not wasted information for a segmenter. See the example shown below:

Types of ready-made meals	person A	person B	person C
Tinned or fresh Soup	x	x	
Pre-prepared Sandwiches			x
Salad box		x	
Italian/ Pasta ready meals	x		x
Pizza	x		x
Chinese / Thai ready meals	x		
British or American ready meals			x
Indian/curry ready meals	x		
Spanish /Mediterranean ready meals	x	x	
Mexican ready meals	x		
Vegetarian ready meals			
Heathier choice ready meals		x	
other types of pre-prepared meals	x		

In this example the segmenter learns that person A consumes a much wider variety of readymade meals than persons B or C, person B is more heath conscious in their choice of food than person A or C. Person C does not seem to like Asian food.

This is a thought worth bearing in mind when designing segmentations. Questions that can do two jobs at once are the most valuable types of question to ask.

In fact, if the preliminary question is asked in the right way about the types of ready meal consumed you might not have to ask the second question about the overall volume at all as this can be inferred from the first.

Predictive questions
It is a lot easier to think back to recent behaviour than estimate general behaviour but in some cases for one of purchases or products with long purchase cycles or where we have no past behaviour to base it on you can switch to a predictive approach. What do you think ….

How much did you spend last time on buying your car?

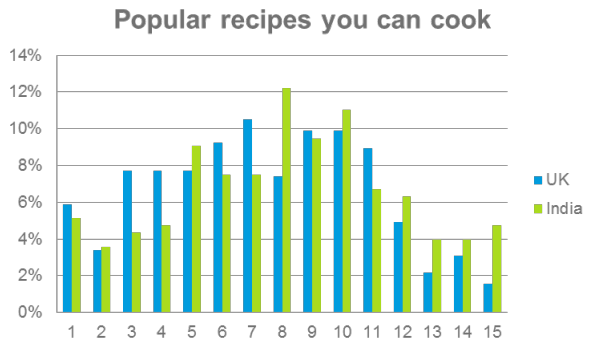
→ *Predict how much you will spend on purchasing your next car?*

These can be made more interesting speculating about what you might do in the future

Test style question
Simple tests are very useful weapons to use in segmentation research to determine how knowledgeable and interested a person is about a topic.

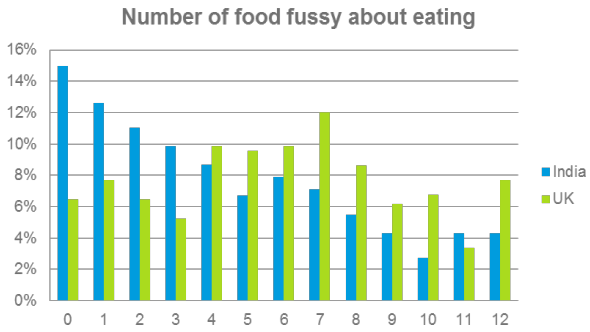
In our experiment we used a range of simple tests to understand how knowledgeable, confident & adventurous people were about cooking. We asked them how likely they were to need to use a recipe book to cook a cross section of popular meals with the option to say they would not feel confident of cooking or would not want to cook. We included some less familiar and more difficult recipes to find out how adventurous a cook they were (figure 20).

Figure 20.



We challenge them in a quick fire round of questions to say if they would eat or not like to eat a range of the top foods people said they were most fussy about eating in each country (taken from our preliminary qual research). We use this to behaviourally measure how fussy they were (figure 21).

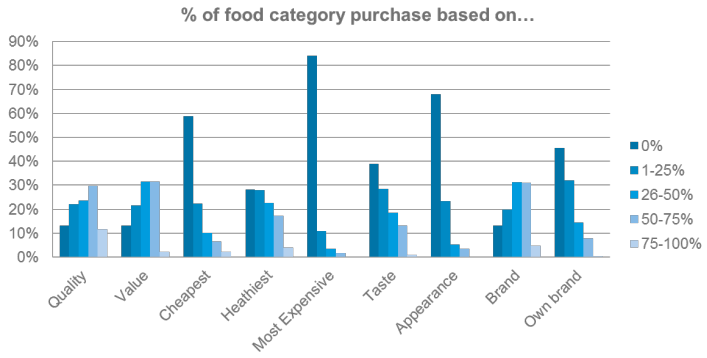
Figure 21.



We also use a similar approach to understand some shopping decisions. Did they have a favourite brand or product when buying a variety of different categories of food? And if not how did they choose, did they pick

the cheapest, the healthiest, the best quality etc… , You can then assess how often they select certain choices across the range of categories (figure 22).

Figure 22.



Max-diff approach

The alternative to these simple test style questions is a max-diff approach where you get people to decide what they would be most or least likely to do. For example, you could present sets of dishes and ask people to pick those they would be most and least likely to cook, eat, buy.

We did not test this question format in this survey as it's something we have extensively experimented in previous experiments and the value of using Max diff scales in segmentation has already been very eloquently argued by Steve Cohen in some seminal ESOMAR papers, e.g. *Measuring preference for product benefits across countries Overcoming scale usage bias with maximum difference scaling,* Cohen & Neira, ESOMAR 2003.

Quiz style questions

You can go a step further and create a proper quiz style challenge where you ask people questions that might test their knowledge of a subject and give them feedback on whether they are right or wrong. Again we have not tested this approach in our segmentation research experiment. This approach is particularly effective for assessing brand or advertising awareness, for example, and for measuring their real knowledge about a category. For example, *Which of these brands of breakfast cereal has the lowest sugar content? Brand A, B or C*

Considering which question style to choose

What hopefully we have demonstrated here is that there are a wide range of different ways of asking segmentation style questions and it begs the questions of how do you choose which ones to use?

Choosing the question format does require some experience. Some questions naturally lend themselves to one format or another. We would recommend where possible to use as many behavioural measures as you can as these provide the most reliable of data but there are some constructs and attitudes that are simply too abstract to ask as anything other than an agreement scale, especially regarding people's personal feeling. It is also fine to maintain a few agreement scale questions in your survey, so long as it's not dominated by them.

There are a lot of efficiency considerations. You have to ask sometimes if it is worth asking this question in a more elaborate way to gain a more accurate answer or is that going to waste a lot of survey time? You can be pragmatic about the choice and mix of questions. The aim as much as anything is to avoid too much repetition of one format of question to maintain the interest of the respondents.

It is all too easy in fact to pre judge the best way to ask each question and so it is our recommendation that the next step of the process is to conduct a pilot survey to test different approaches.

Stage 4: Piloting

You will probably, as a result of the process you have gone though so far, have generated a great big long list of questions and some of them you are unsure the best way to ask them. You will also have probably far more questions than you will ever be able to fit reasonably into a practical length survey.

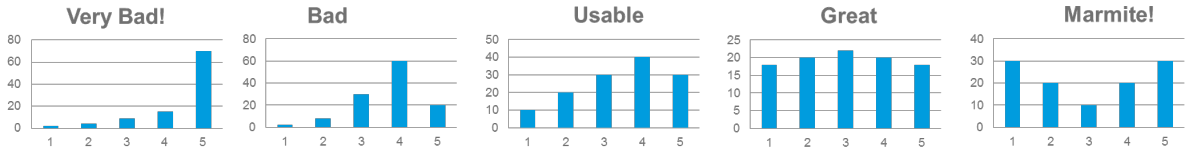
The role of piloting is therefore to identify first the most efficient questions to use and then secondly the most efficient ways of asking these questions. This process is a fairly straightforward one but it is best done with a stop and start, flexible approach of test and refine.

Looking at the diversity of the data it becomes clear which statements divide opinions and which result in strong agreement or disagreement. If one is looking to build a multi-level complicated algorithmic segmentation then one could be looking for the minority who go against the general trend to set against different data points. However generally speaking one does not gain much insight from such data points.

Finding the Marmite* questions

Instead the best questions are the divisive ones that split the audience clearly and relatively evenly and the most useful ones of all are those that completely divide the population into two clear distinct groups – these Jason Brownlee from Colourtext refers to as the “Marmite questions*”. If you can find one or two of these in your pilot survey you will be very lucky!

Figure 23.



These ‘Marmite’ questions can be supplemented with a selection of questions which aren’t as divisive to look for the few individuals with outlying opinions. There are no hard and fast rules to this but we would recommend that at least two thirds of your questions be ones which have a relatively even distribution across the scale. However these should not form the basis of your approach as the key driver for a segmentation is clearly segment. Overreliance on multiple small differences to build a group is an indicator of a weak segment.

Correlation analysis can also be used to identify the questions which overlap. Obviously causal relationships may be interesting to understand and identify the individuals who buck the casual trend. However generally, as with non-divisive statements, too many causal questions ultimately means you are asking the same thing in a different manner, and hence its best to avoid questions which correlate strongly except in a few specific circumstances.

*** What is Marmite?**

To those of you who are not familiar with Marmite, this is a British spread made of yeast extract that notoriously divides opinion in the UK, half the population loves the taste and the other half hates it.

Stage 5: Crafting the design of your survey

So as a result of your pilot you whittled down which question you want to ask and how you want to ask them, and the next step is crafting the design of the survey.

First of all you are going to need to think how you are going to pitch your survey to respondents.

The quality of feedback you get from a segmentation survey is heavily dependent on the attitude of the respondents completing the survey. It’s really important to get them into the right frame of mind, you have to win their trust that this survey is going to be an interesting and fulfilling experience.

We recommend investing in the design of a visually engaging intro screen. Figure 24 shows the simple intro page we created for our segmentation survey experiment as an example.

Figure 24.

As a way of engaging respondents, an approach we like to use is to present a diverse series of candid visual "memes" reflecting the range of thoughts and feelings people have about the segmentation topic and ask them to think about how much they identify or not with each one.

The aim of this process is to make it clear that people have all sorts of different and often conflicting thoughts and feelings and the survey is a safe environment to express your honest feelings and behaviour.

Figure 25 provides examples we created for our cooking segmentation experiment.

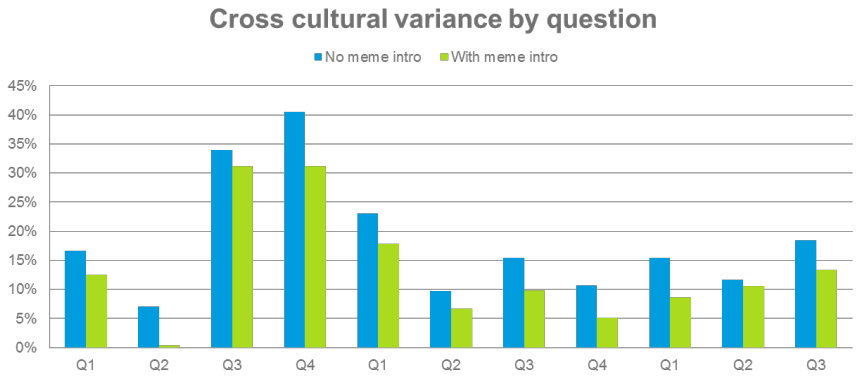
Figure 25.

This might seem a superfluous addition to the survey but we can demonstrate this approach can have a measurable impact on the quality of feedback we get from the rest of the segmentation survey.

In this experiment we found those groups exposed to these intro meme questions invested an average of 10% more time in thinking about their answers to subsequent questions. It resulted in more candid and

honest responses and helped to consistently reduce the levels of cultural variance by an average of 30% (figure 26).

Figure 26.



Story boarding

One of things we most focus on when designing segmentation studies is building them around a clear narrative structure. What respondents find most frustrating is being presented with a jumbled mass of questions with no sense of order and what is coming next.

There are lots of readymade survey narratives that can be used. The journey to purchase, for example, is a very effective one. You start out asking them how they would go about buying a product, what they think about, what advice they take and information they would use. Then ask them to think about what they have purchased in the past and why and what they would want differently in the future.

Placing products and services on trial or asking people how a product fits into their life or creating the perfect product are all good narrative themes.

At the start of the survey you introduce the narrative theme, explain what is entailed, and at each section of the survey try to explain why you are asking the questions within the context of the narrative.

For example, imagine the theme for this experimental survey was “how does cooking fit into your life”. You might introduce the survey with explanation something like this…

“In this survey we want to understand how cooking fits into your life, your feelings and attitudes towards cooking. To help us understand this we are interested in knowing what you like to cook, how often you cook and who for”.

Maintaining respondent interest

Having published several previous papers on how to engage respondents we will not dwell upon the technique you can employ to make the survey more engaging. One example is “Can gaming techniques cross continents”, Puleston & Rintoul, ESOMAR 2012.

As a general rule, to maintain respondent’s interest we recommend using a good mixture of different question formats and styles. Try to avoid asking repetitive banks of questions in the same way. We recommend breaking up the flow with what could be termed “quick fire” rounds, where you set tests that you encourage respondents to answer spontaneously. Empower them to make decisions, cast their opinions and encourage them to think.

There is no law to say you can’t make it fun! And where possible this is encouraged.

Providing feedback

Give feedback where possible and there are various ways this can be done. We have found that telling respondents how they are getting on in a survey or feeding related information to them is a great way of increasing engagement. In side by side testing we have seen the addition of feedback telling people how their answers compare to others can lead to people taking 20% longer to answer questions, and 15% higher standard deviation in the answers they give.

The point that we overlook is that respondents take surveys mostly to express their opinions and to see how they can have an impact. Feedback not only provides something fun and more conversational, but also feeds into this desire.

The beauty of segmentation is that we have an in-built feedback mechanism, namely the segments themselves. Simply by telling people which segment they fall into at the end of the study we have turned essentially a rather dry task of self-observation into a personality test.

Obviously some care has to be taken to deliver the feedback in a way that will not bias the data or offend the respondent but this is very easily done. This is an approach which we have standardised and implement on many segmentation studies with great success.

Timing and costs of delivery

Some readers of this paper might well have some valid concerns about the time it would take to set up a segmentation survey in this way as the procedure involves at least two preliminary stages of discovery and piloting research and this could also increase the overall cost of the project.

In our experience, having a formal framework to develop the survey and by doing proper discovery research an piloting of the questions and having concrete evidence as to how a particular question is being answered removes a lot of need to speculate and reduces the toing and froing and debating over the content of the questionnaire, so surveys can be signed off in much fewer drafts.

Working in the dark, without this preliminary research, the process of writing the questionnaire can become an extremely drawn out process. It can result in a committee of decision makers all suggesting lots of micro-refinements to questions based on speculation about what might be the best approach and involve multiple drafts of the survey. Ten or even 20 draft versions of a questionnaire is not uncommon in these circumstances, adding weeks to the overall development process.

So we would argue these formal procedures can reduce overall set up time. Much of the extra work can be conducted concurrently. The initial research discovery phase, using a quali-quant online approach, takes less than a week to implement and can be conducted at the same time with the early questionnaire scoping process.

The second stage piloting process adds approximately one week to the set up process but you can gain this time as it reduces the uncertainty about which questions to ask and the best way to ask them and so sign off procedures are usually much more efficient.

Adopting an agile iterative approach, much of the time taken to set up and field a segmentation study is in the sign off procedures for the design of the questionnaire and can be reduced.

Cost considerations

In terms of extra cost of the preliminary research, it will require potentially 200-400 more units of sample. The cost of this has to be weigh against efficiency gains in the quality of the survey and data you gather and the overall number of completes that will be gathered over the whole lifecycle of the segmentation project (often several years).

Piloting can significantly improve the overall quality of a segmentation by allowing you to test out a broader range of initial question topics and then focus more efficiently on the topics and questions that clearly differentiate opinion. This provides an important insurance against the serious risk of failing to cover key aspects of consumer behaviour and attitudes within a survey. All researchers and clients know how bad it can feel when it becomes clear towards the end of an analysis that something of crucial importance in consumer decision making has been omitted from an expensive research study. Failing to cover topics of such importance within a segmentation exercise will seriously degrade the final value of a project - our approach to segmentation piloting removes this expensive, and often career-limiting, risk.

Summary and timing delivery estimates

The basic set up process when conducted efficiently takes approximately six weeks.

Week 1: Discovery phase

1. A preliminary quali-quant style piece of research to map out the decision making drivers
2. Augment this by identifying the unconscious decision making drivers

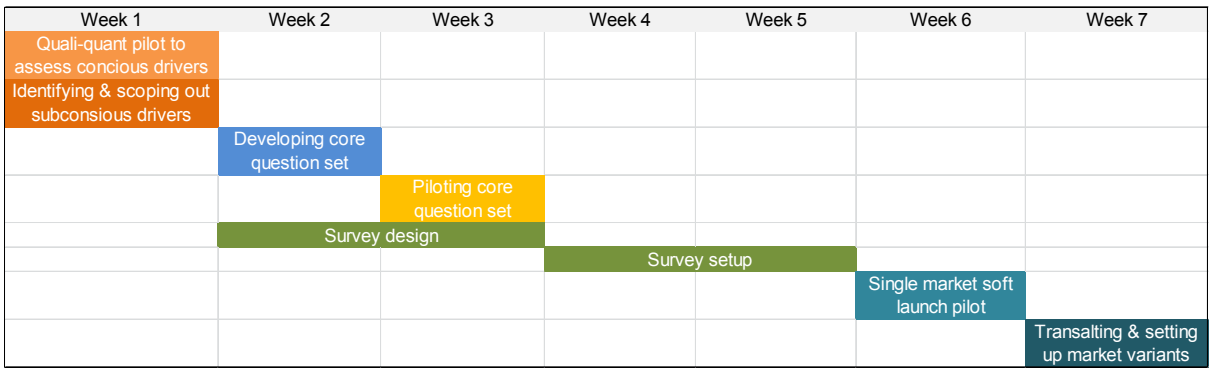
Week 2 & 3 Scoping & piloting of the core questions

3. Establish a list of core questions
4. Carefully consider the best way to ask each one to minimise cultural biases, we recommend:
    - Including a calibration process
    - Using bespoke range anchoring
    - Using comparative scales
    - Behavioural based questions that are manifestations of attitudes
    - Focusing on recent behaviour and using memory prompts
    - Test style/max-diff and quiz based question formats
5. Once you have established the ways you want to test asking each question, you then set up and conduct a basic pilot:
    - To identify the most efficient questions
    - To test the best way to ask questions
    - To streamline the survey

Week 4 - 6 Survey design phase + 1 day per market variant

6. Finally designing the survey:
    - Think about how to pitch the survey to respondents
    - We recommend the use of “memes” to engage respondents & encourage them to open up and give more truthful responses
    - Build the survey around a strong narrative structure
    - Try to introduce some elements of fun
    - Where possible provide respondents with feedback
7. Once signing off the survey we recommend a single market soft launch pilot to around 100 respondents to check the survey is running efficiently
8. Then after this the survey will need to be translated & local market variants created

Figure 27.



Note these core timings do not take account of the vagaries of sign off procedures which will vary project by project & by client.

Once you have done all this your survey is ready to field.

You then have the challenge of analysing the data, developing the segmentation, but the techniques involved in doing this are for enough for a whole paper in itself so we will leave it here….

References

Survivorship Bias: https://en.wikipedia.org/wiki/Survivorship_bias

Segmentation: https://en.wikipedia.org/wiki/Market_segmentation

History of segmentation: https://datablog.visualdna.com/2015/01/27/history-of-segmentation/

Patterson M ESOMAR (2016) Datactionable! Actionable segmentation: fusion of customer and survey data

Cohen S, Markowitz P ESOMAR (2002) Renewing market segmentation, some new tools to correct old problems

Cohen S & Neira ESOMAR (2003) Measuring preference for product benefits across countries Overcoming scale usage bias with maximum difference scaling.

Montemayor MC ESOMAR (2003) Segmentation beyond cultural boundaries

Morris L, Hedger G, Schmolze R ESOMAR (2004) Forever Jung. Identifying consumer archetypes to help guide communications strategy

Oustrup MK, Jeanes M ESOMAR (2005) Mood Consumption® Theory. A human-focused marketing tool

MacLachlan D, Mulhern M ESOMAR (2004) Segment Optimization. An empirical comparison

Desai S, Costella T, De Vos K, Bond J, O'Brien T , Davis S, ESOMAR (2009) Mastering Insights That Pay. Tying the knot with insights: the ingredients of a successful segmentation

Counter view on the value of segmentation:

Bergström F, Ones J Stockholm School of Economics (2013) The Lacking Relevance of Market Segmentation. A cross-category study using the Dirichlet model

Cross cultural segmentation studies survey design considerations:

Pettit A, Courtright M, Pashupati K, ESOMAR (2013) Multimode, Global Scale Usage. Understanding respondent scale usage across borders and devices

Puleston J, Cashman A (2015) It's not business, it's personal: the power of personal engagement in Business to Business research. MRS

Puleston, J (2014) Bonsai Survey Design ESOMAR Asia

Puleston,J Eggers,M (2012) Dimensions of online survey data quality ESOMAR congress

Rintoul D & Puleston,J (2012) Can Gaming Techniques Cross Cultures ESOMAR Asia

Malinoff, B & Puleston, J. (2011) How far is too far? Traditional, flash and gamification interfaces, and implications for the future of market research online survey design ESOMAR 3D Miami.

Puleston, J. & Sleep, D. (2011) The game experiments: Researching how game techniques can be used to improve the quality of feedback from online

Sleep, D. & Puleston, J. (2009). Panel quality: Leveraging interactive techniques to engage online respondents. ARF Congress, http://www.gmi-mr.com/solutions/resources.php.

Puleston, J and Sleep, D (2008) Measuring the value of respondent engagement: Innovative techniques to improve panel quality. ESOMAR Online Dublin

The Authors

Jon Puleston is Vice President of Innovation, Lightspeed Research, United Kingdom.

Alexander Wheatley is Innovation Researcher, Lightspeed Research, United Kingdom.

Jason Brownlee is Founder and CEO, Colourtext Ltd., United Kingdom.

.

PART 5
Engagement Discoveries in Media and Retail

Indian TV Industry

How the stars in the sky changed

Romil Ramgarhia • Bipin Mundhwa • Rucha Deshmukh

Introduction

Audiences no longer consume TV unilaterally. While current technology captures the audiences who are watching TV, it does not take into account their actual behaviour whilst consuming it.

This paper attempts to answer the various multitasking/activities that audiences from metro cities are doing while consuming TV content. What does divided and undivided attention on TV mean in the context of product and brand recall?

TV consumption with multitasking/divided attention can be further probed to understand whether the time is right for audience measurement agencies to enter into behaviour based measurement (Real Audience) in addition to the current head count based measurement.

Knowing India and challenges

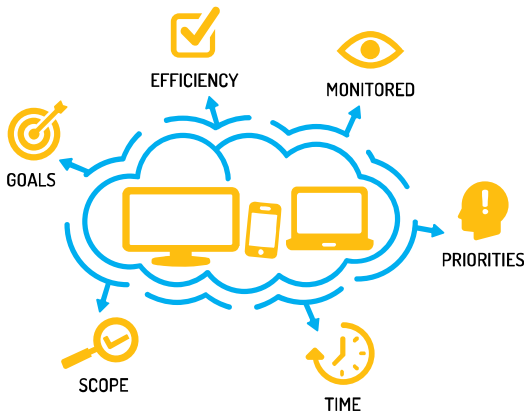
India is not just a country but a culture of coexistence of multiple ways of life. Besides being the world's largest democracy, 17.5% of the world's population resides in India. If we were to look for a point of comparison to understand how big the population of India is, we would have to look at not another country but entire continents - Africa has 15% of the world's population and Europe has 10% of world's population.

India has a population exceeding 1.21 billion which is spread across 29 states and 7 union territories.

There are about 122 different languages with various dialects. Of these, 22 languages are recognized by government of India as an official language. 29 languages have more than a million native speakers and 60 languages have more than 100 thousand speakers.

We say India is a country, but within India there are many smaller countries. The diversities in the context of audience measurement come in form of multi genre content with multilinguistic TV channels within each. Additionally, new age media options, like social media, show correlation with television.

Figure 3.



Other major diversities that add complexities to audience measurement are the multiplicity of languages and dialects with wide geographical spread, multi-cultural and multi-religion society leading to distinct festivals, food habits and preferences.

Figure 2.

In addition to that, proliferation of media is increasing over the last few decades and it is growing exponentially. For example, in 1980 there were just five main stream media options and today we have multiplicity of options that is still growing and providing the advertisers with a platter of options.

If we combine all these factors, such as population, multi linguistic, multicultural society with multitude of media options, we can see why India is challenging market for audience measurement.

The need deficiency for new TV audience measurement

Let us understand the need for a new TV audience measurement in India. In other words, why was there a need to replace the incumbent and introduce a new audience measurement system.

Firstly, despite a huge population and diversity, in terms of religion, culture and spoken languages, the incumbent was measuring only urban India and not rural India.

Secondly, there were no stake holders from the TV industry as part of this incumbent. On the contrary, the incumbent had roots in a global media agency group creating conflict of interest.

In the past, there were many issues and data anomalies reported by broadcasters, agencies and advertisers and they continued growing in numbers rather than getting resolved. Eventually, a legal suit was filed by one of the broadcasters raising questions on vigilance and malpractices.

In addition, the Ministry of Information and Broadcasting (Government of India) came up with guidelines that specified anyone measuring TV audience should have a minimum of 20,000 sample households to start with and it should gradually increase to a 50,000 sample. The incumbent had approximately 10,000 sample households only with no plans of increasing them in the future.

The guideline specifically stressed that no private player with conflict of interest can fully own the TV audience measurement. These were few of the major areas where incumbent failed to fulfil the requirement.

In a nutshell, we can say that the need was acutely felt by all the stakeholders for a new TV audience measurement system in India.

The challenges were tackled

Let us understand how the challenges discussed above were tackled. To start with, BARC India set up a panel with 22,000 households in keeping with the guidelines by Ministry of Information and Broadcasting. Unlike the incumbent, BARC India had no conflict of interest as this was a joint industry body of broadcasters, media agencies and advertisers. BARC India, as a joint body, was set up with fair and transparent policy as its founding pillar. It is open for audit for the betterment of industry.

When BARC India was set up, we complied not only in terms of policy guidelines and objectives of fair play but also introduced several technological breakthroughs. BARC India started with watermarking technology in India for tracking and reporting of channels. This ensured data hygiene at ground level.

Over and above this there was introduction of a new socio-economic grid for segmenting the audiences which replaced the old socio-economic grid. The new socio-economic grid is based on more practical variables for ascertaining the social status of audiences. The new variables under new socio-economic grid are:

- Education of chief wage earner
- Number of household durables

Figure 3.

This can define the social status and purchasing power of the family members in the household.

The milestone achievement for BARC India was successfully launching Rural India viewership. This was the first time in India that Rural viewership was being measured and reported to the TV audience measurement stake holders namely advertisers, broadcasters and media agencies.

Figure 4.

Let us understand the technological process of audience measurement used by BARC India. The broadcasters watermark the channel ID, date and time of the content. When this content is being viewed in the household the Bar-o-Meter captures it along with profile of the audience watching it in the household. This generates the raw data of viewership at the local polling station with the help of the GSM partnership. The same data is then weighted with the help of an establishment study of the audience and demographics. The entire data is then mapped with content played (Program tagging) across various channels by time and date. This is reported in the final output in the BMW software.

The technology by BARC India was much advanced in comparison to the incumbent mainly due to the adoption of watermarked technology of channel. In a nutshell, a watermark is a unique audio signal assigned to each channel that is not audible to humans. It is audible to the measuring device only. This technology is almost two generations ahead of the incumbent. It was adopted by France in 2007 and by United States in 2013. The key benefits of watermarking are, firstly, time shifted viewing measurement. Therefore, viewers who are delaying the viewing of their choice of content for later in the day can also be captured. Secondly, it captures the same content on two or more different channels. This is called simulcast measurement. Finally, it is also a future ready technology as it is compatible with digital measurement technology at an extremely low cost.

The measuring device is called a **'***Bar-o-Meter'* which is powered by 'Intel Chip' as the processor. It is manufactured in India with almost 1/6th the cost of globally available meters. It uses a third generation OLED display which is a first in the world. Importantly it is mapped with all key telecom operators across India that ensures real time capturing of data.

The entire data capturing and reporting is designed in such a way that it ensures smooth functioning whilst ensuring no one can influence or overpower anyone else. There are namely 12 major processes that BARC India under goes with 30+ vendor partners who are specializing in their respective fields. The entire set up is done in a way that secures the information integrity with a strict cost consideration.

BARC India has come up with a completely fresh and user-friendly software that provides wide options to users in partnership with Markdata, one of the world's leading audience measurement system provider. The name of the software is 'BMW' − that is 'BARC India Media Workstation'.

The financial challenge was tough to overcome as BARC India was nonexistent in any media or audience measurement business. BARC India was formed as a not for profit organization like any other charity or non-government organization. BARC India did not borrow money from any private lenders but raised money from banks with the stakeholders standing as guarantors for it. There was no financial burden on any particular stakeholder to recover the money. The profit arising out of BARC India is not supposed to be shared with any stakeholders but to be utilized for the purpose of audience measurement only.

Figure 5.

To sum up the stakeholders in this joint industry body, there are 111 advertising agencies that form the Advertising Agencies Association India, 53 broadcast companies with 800+ TV channels that forms the Indian Broadcast Federation and finally a 155 advertiser group that forms The Indian society of Advertisers.

Relevance

Winning the trust with uniqueness and new initiatives

Let us understand the future plans of BARC India. Very soon the panel would be expanded to 30,000+ households. This will provide a wider and deeper spread for robust data. In a country, as big as India, a wider and bigger spread of the sample is very critical for accuracy.

BARC India is also going to come up with TV+Digital measurement for India going forward. For this, the audio-visual content will be measured across media platforms. This would be for the first time in the world where a unified measure of Impressions across media platforms for audio-video content will be reported. It will add tremendous authenticity to data as it will be a third-party measurement system (neutral party) and not from publishers.

BARC India presents another unique product called 'Video Asset Link Identification' (VAL ID). This is a 27-digit unique watermark assigned to each Video Advertising Asset across TV and Digital media linked with detailed Advertising Metadata. It helps create a single identity for exchange of information across agencies and broadcasters. It would greatly help with spot reconciliation and payment settlement due to unique identification.

Figure 6.

BARC India has started BARC Media Lab which is called “THiNK’. This media lab is responsible for data mining for adaptive insights with trend spotting and causal relationship of data variables. One of the core offerings of BARC India Media Lab are the THiNK Newsletters which provide actionable insights for advertisers, agencies and broadcasters. Additionally, THiNK Media lab is also doing social media listening and benchmarking to understand the impact of TV content on social media and vice versa.

BARC India plans to come up with a product for forecasting of audience impressions to help media planners and buyers take informed decisions and make media plans.

BARC India has entered into a joint venture with TAM (previous TV audience Measurement Company) to take the benefits of fusion of data of past and current measurement. This joint venture will help BARC India to enhance current offerings and capabilities with a wider sample distribution.

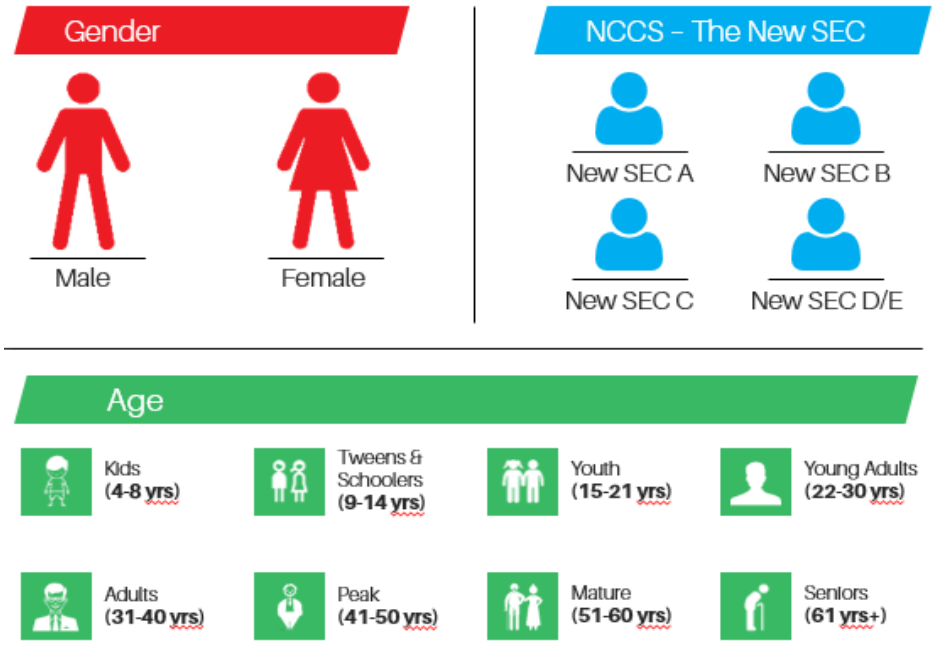
BARC India has also introduced a certification test to assess an individual’s knowledge of BMS (TV Planning/rating software). This certificate is aimed at all BMW users – agencies, broadcasters and advertisers. This certification could help add quantifiable credentials to individuals and corporates.

Figure 6.

What is revealed by the TV panel?

BARC India finds it imperative to understand the behavioural aspect of its TV panel households as it employs one of the largest TV panel system. The panel homes are spread across India and it unlocks a wealth of information that is being captured and reported in BARC India proprietory software BMW.

Figure 7. Reporting parameters give a more granular view of the audience profile



Key Audience Variables	Markets
Impressions	Urban and Rural
ATS	
Share	17 states
Reach	
Coverage	3 Urban pop strata – Mega cities, 1-7.5 million, <1 Million
OTS	

Along with the audience measurement, it is important to measure and understand the qualitative aspect of viewing patterns and behaviour that can be mapped with quantitative audience measurement.

Background and literature review

The audiences no longer consume TV unilaterally. Today's fast moving life in Indian metro cities coupled with the proliferation of media platforms, as well as spread of genres and content type, makes audiences consume TV with divided attention or while multitasking.

While the bar-o-meter captures the audiences who are watching the TV it does not take into account their actual behaviour whilst consuming TV. However, it is important to capture this behaviour of audiences as it potentially has an impact on the product and brand ad message receptivity.

According to Green (2010), there are number of factors that impact the recall of ad messages on TV namely, ad position, number of other commercials around it, length of the ad, etc. One of the large study by Cable Advertising Bureau in the United States suggests that early position in the break, shorter breaks and fewer brands competing in the same product group impact the aAd message recall.

However, one of the potential factors that also impact ad message recall with respect to product and brand is the level of engagement/attention with the TV content. The described study on TV panel households of BARC India is focusing on the engagement/attention level that can be classified as divided and undivided attention while consuming the TV content. According to Cannon and Riordan (1994), effective reach frequency is more of a theoritical approach. The media vehicle and actual ad exposure are distinct and they cannot be merged or assumed as having halo effect. The study discussed here does not attribute ad message recall with respect to media vehicle exposure and frequency of ad message. The study is

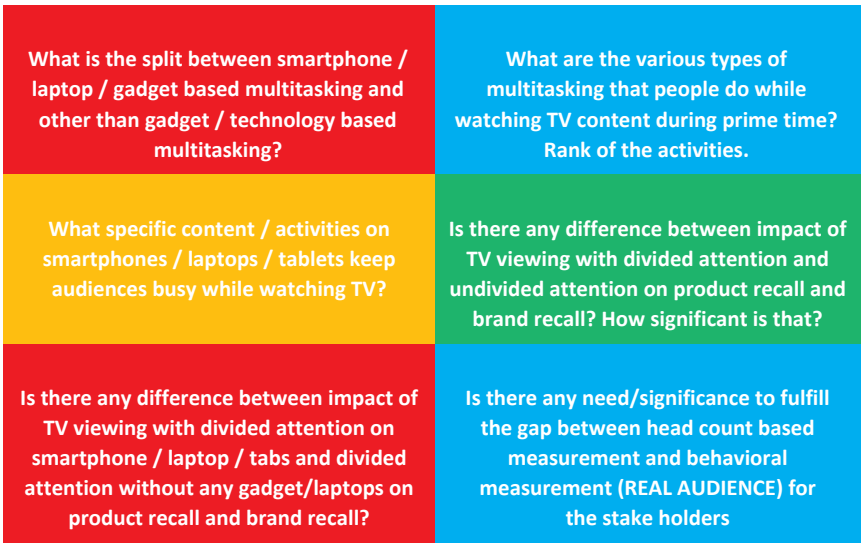
discussing the behavioural factors of TV viewing habits that most of the contemporary papers inadequately address.

According to Carruthers (2016), there are two types of engagement/ attention, namely selective attention and alternating attention. It says that it is about choice and people spend more time on the screen that attracts them. The paper by Carruthers focuses on engagement and not its impact or relationship with product and brand recall.

Audience behaviour study

The questions that we are attempting to answer with our research experiments on metro audiences in India are summarized in figure 8.

Figure 8.



Methodology

The study was conducted on panel households from three mega metros of India (Mumbai, Delhi and Bengaluru) with participation of 300 households across NCCS ABC 15-40 years. The above findings are only for the prime time band 1900-2400 on YVR (Yesterday Viewership Recall) during Jan-Feb 2017.

In order to analyse the audiences, they are classified as audiences with *Divided Attention* and audiences with *Undivided attention*.

Definition

- *Divided Attention* – Audience who is either secondary viewer or watch TV along with some kind of parallel activities, either gadget based or non-gadget based, is considered as audience with divided attention.
- *Undivided Attention* – Audience that watch TV as a primary viewer, without doing any parallel activities/disturbance is considered as audience with Undivided attention. (See figure 9.)

It can be seen that TV content consumption in metro cities of India is usually multi engagement (Divided attention). Most of the audiences consume TV content when they are busy with gadgets/smart phone/tablets during prime time of TV.

Figure 9. Divided vs. undivided attention (watched TV for at least 60 minutes)

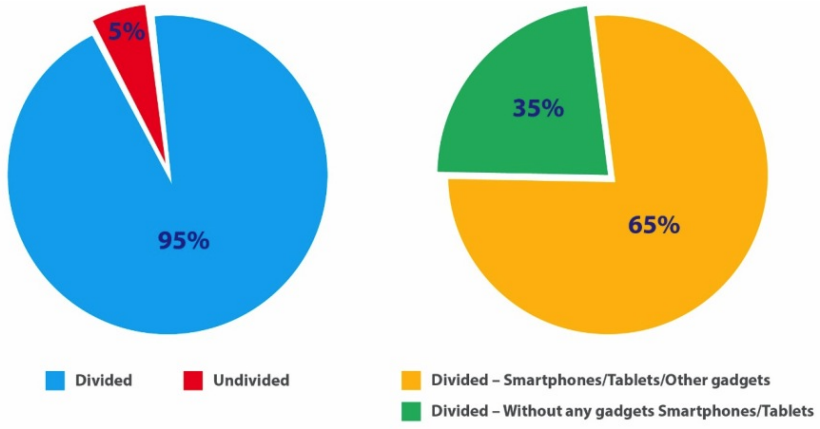


Table 1. Various activities on smart phones / tablets / laptops while watching TV

Activities	Rank in Ascending Order
WhatsApp	1
Social Media	2
YouTube/Video	3
News app	4
E-mail	5
Shopping	6
Music app	7
Others	8

Table 2. Various activities / engagement while watching TV other than smart phones / tablets / laptops

Activities/Engagement	Rank in Ascending Order
Secondary viewer	1
Non-gadget based in same room (Eating, cutting veg, Cooking, Reading, Talking)	2
Switching between rooms	3

Among the top activities on smartphones / tablets / laptops while watching TV with divided attention are chatting based Apps and Social media. These can be safely considered as a very regular and frequently used activities even during prime time of the TV. Using Shopping and Music apps are among the least done activities while consuming TV content.

Table 3. Various activities while watching TV by number of product categories, brand recall, brand consideration and brand preference

Watching TV and was busy with smartphone/laptops/ tablets on…	# of Product categories recalled (Avg.)	Brand Mention (Avg.)	Brand Consideration (Avg.)	Brand Preference / Intention to purchase (Avg.)
Social media	6	9	4	2
Gmail/Email	3	6	3	0
News app	4	5	3	0
Music App	2	4	1	0
YouTube/ Other Video content	2	5	2	0
WhatsApp / Chatting apps	5	8	4	1
Shopping App	3	5	1	1
Any other App	2	4	2	0

The number of product categories, seen on TV, recalled with divided attention on smartphones / tablets / Laptops is relatively higher for Social media and App based chatting. This could be because these activities consume time but don't take up as much attention as other niche activities such as shopping app and other apps. This is evident from the lower number of product recall done by audiences who were busy with divided attention on not so regular/niche apps.

Interestingly, in path-to-purchase funnel the audiences who were busy with social media and WhatsApp happen to recall more number of brands and it goes down to the level of brand preference or intention to purchase. The less regularly used apps may be stressing little more attention from audiences as number of product and brand recall is relatively less as can be seen from the table above.

Social media/chatting apps have become the part of our normal life. Multitasking, especially whilst busy on social media is seen as a normal behaviour and less stressful. We have gradually become accustomed to social media chat and also watch/ listen to TV or talk simultaneously.

On the other hand, Music app and email reading requires more attention and can be relatively stressful for several reasons. These could be reading finer points in email, thinking of what to reply to email, how to reply with right vocabulary, changing or searching the songs, trying to skip advertisements and play next music items. In addition, it can be said that email and listening music seem to be relatively less frequent in usage compared to social media chat. These all collectively appear to be the potential reasons for social media/chatting apps register a better recall and brand consideration in comparison to Gmail and Music apps.

Table 4.

Watching TV without attention on smartphone/laptops/ tablets	# of Product categories recalled (Avg.)	Brand Mention (Avg.)	Brand Consideration (Avg.)	Brand Preference / Intention to purchase (Avg.)
I was watching because someone else in the room wanted to watch the TV	4	7	4	1
I was switching between rooms and was not attentive at all for TV set	2	7	2	1
I was watching TV without disturbance for most of the time	6	9	4	3

TV audiences, during prime time, recall the product and brand much better when they are watching TV without any disturbance. Switching between rooms while watching TV (pulse level of TV content consumption) is the least effective as far as product and brand recall are concern.

Table 5. TV viewing with divided and undivided attention – Type of product categories recalled (need to look at for variety)

Rank	Divided	Undivided
1	Soap	Soap
2	Shampoo	Shampoo
3	Toothpaste	Toothpaste
4	Oil	Banking
5	Chocolates	Cream
6	Biscuit	Paints
7	Washing Powder	Mobile handset

It is evident that audiences with undivided attention tend to remember a variety of product categories whereas audiences with divided attention skew towards product categories from the FMCG sector, that is categories with a high frequency of advertising on TV.

It is important to note that FMCG as a sector predominantly enjoys the highest share of commercial time on TV and relatively stays above other categories in terms of impressions. Therefore, chances are brighter that an audience will get a higher exposure to FMCG brands on TV. This could be the reason why FMCG as a sector is easy to recall.

As far as brand consideration or preference is concern, the biggest strength of FMCG itself becomes its weakness as it is highly cluttered with multiple brands spread across dayparts. This tends to impact the decision makers brand consideration and brand preference stage in the path to purchase.

Whereas Non FMCG product categories are seasonal and highly need driven in terms of brand consideration and preference. This makes non FMCG product categories relatively easy ones to get shortlisted for consideration and preference.

In the given context, it is quite easy to have FMCG categories as top of mind recall as they are present on TV space throughout the year across dayparts. Due to this the divided attention of audiences does not have impact on recall. However, undivided attention will give more variety in the product categories list as audiences are attentive and may recall product categories other than FMCG as well.

Table 6. Perceptual understanding of TV viewing and TV ad recall

Sr. No.	Statement	Agreement (Top 2 Boxes)
1	I usually remember brand advertisements on TV better when I am not multitasking/doing multiple activities while watching TV.	24%
2	I prefer watching TV alone and doing no other activity simultaneously.	37%
3	I enjoy watching TV along with doing other activities e.g. Apps on Smartphone, Eating, talking on phone or with someone in room, cooking, etc.	53%
4	I remember the content of the programs on TV better when I am not with anyone else or not doing anything else while watching TV.	48%
5	I don't understand the Ad message/advertisement when I am doing multiple things while watching TV	43%
6	Most of the time I watch TV while doing multiple activities at Home e.g. busy on Phone, Apps, Cooking, talking, switching between rooms, etc.	61%
7	I usually watch TV when I am doing other tasks/activities	74%
8	Brand advertisements on TV are interesting / helpful in my purchase decisions	83%
9	Even if I am multitasking, I pay more attention to program content than advertisements	67%
10	Alerts on News, Sports, TV series story, lifestyle, food, travel, etc on my phone makes me switch to the TV in search for the related content.	67%
11	When I see brands/products on TV, I search them online using smartphone, laptop, tab, etc.	78%

In totality, it is safe to say that divided attention while watching TV, even during prime time, is a reality. However, TV as a media platform helps audiences in making purchase decisions. Audiences, don't see divided attention while watching TV as disturbance, on the contrary most of them seem to be enjoying multitasking (divided attention) while watching TV.

TV prompts the audiences to search for the brand online. Similarly, News, Sports and other alerts on phone make audiences switch to TV in search for the related content.

The main idea behind capturing statement 1 is to get pure recall and not decoding of ad message (Creatives). Whereas, statement 5 indicate towards the understanding of ad message or storyline in the TV commercial. It can be said that when audiences are doing multiple things they may recall brand but chances are brighter that they may not understand the storyline in the TV commercial.

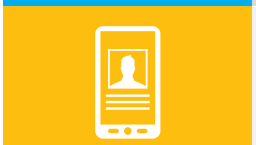
As far as brand recall is concerned, TV definitely plays a role as can be seen in the data/experiment. It is interesting to see what role do platforms other than TV play in path to purchase decision. TV as a media platform still remains the most impactful platform in the Indian market.

Secondly, the statements are capturing the perceptual understanding of TV viewership from the audience's perspective. What they feel about TV as a media platform has been captured as claimed psychographic statements. It may not necessarily add up as being psychographics statements rather it provides broad base to substantiate assumptions and comments about TV as a platform. In measurement science we capture the action like watching TV and not what we feel or perceive. Therefore, these psychographics statements provide base to articulate the comments.

Conclusion

TV could be termed as one of the most effective media platforms as even divided attention does not impact the product and brand recall in metro cities of India.

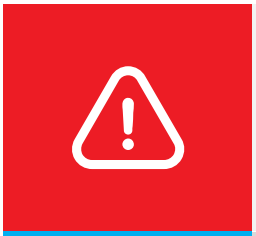
Social media and App based chatting are the top content that are being consumed parallaly on smartphone / tablets / laptops while watching TV. It could safely be termed as undivided attention as it does not impact the final stage of path to purchase i.e. brand preference or intention to purchase.



The apps that are not so regular or not so frequently used may require more attention of audiences. This may potentially impact audiences' product and brand recall in the path to purchase funnel.

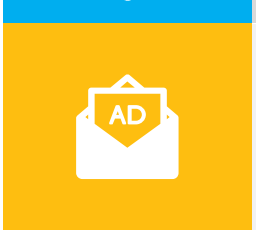
Audiences with undivided attention tend to recall more variety of product categories as compared to audiences with divided attention.

Scope for future research / discussion points



TV consumption with multitasking/divided attention can be further probed to understand whether the time is right for audience measurement agencies to enter into behavior based measurement (REAL AUDIENCE) in addition to the current head count based measurement.

Divided attention can further be probed to assess trust factor of the audiences for advertisements on TV vs Print vs Radio vs Internet and in combination of all these media platforms.



It would be interesting to ascertain the awareness drivers, consideration drivers and preference drivers for various audience segments, when they are exposed to same ad message across media platforms.

The Authors

Romil Ramgarhia is CBO, BARC, India.

Bipin Mundhwa is GM, BARC, India.

Rucha Deshmukh is Senior Manager, BARC, India.

.

Shifting Customer Engagement in the Multichannel Retail Era

Understanding changing behaviours, delivering engagement and measuring ROC (Return on Customer)

Teeradet Dumrongbhalasitr • Vijay Balaji M

Introduction

Consumers are making significant changes in the way they shop for groceries and this has important implications for brands and retailers. In grocery retail, across the globe, many brands now see one-fifth to almost half their sales from online purchases. However, not all markets adopt online grocery at the same pace. Do you know that 99.9% of Thailand customers shop for products in-store? 55.8% of the Thailand's population is connected to the internet. There is a need to form a bridge between digital technology and grocery purchase and the answer is not 'online' at the moment for certain markets.

This study presents the findings from a global study of consumer shopping behaviours carried out by dunnhumby and how location based mobile offers deliver impact on customers, as well as how to measure customer engagement and how it translates into sales and profitability of a business.

Study objective

a) Understand the change of customer behaviour in the multichannel era
b) Deliver new customer engagement strategies to trigger customer purchase
c) Measure the campaign to deliver ROC (Return On Customer)

Context

Online and multichannel shopping has become one of the most discussed trends of our time. Discussions on online and multichannel shopping in the grocery space gravitate quickly to the extremes. Many will make profound and hyperbolic statements about how online will fundamentally change grocery shopping, while others are more sceptical about widespread consumer adoption of online shopping in this category. The fact is, consumers are making significant changes in the way they shop for groceries across online and in-store channels and this has important implications for brands and retailers. In the grocery retail category, many brands now see one-fifth to almost half their sales from online purchases. Growth is expected in all aspects of the market.

Figure 1. Percentage of shoppers by type and index of multichannel sales

MARKET CATEGORY	IN-STORE ONLY	ONLINE ONLY	MULTICHANNEL SHOPPERS	MULTICHANNEL SALES INDEX	COUNTRIES CLASSIFIED IN THIS CATEGORY
ESTABLISHED	83.5%	0.2%	16.3%	130	UK, South Korea, France
EMERGING	96.8%	0.1%	3.2%	127	China, Slovakia, Czech Republic, Poland, Ireland, USA, Japan, South Africa
NASCENT	97.5%	0.1%	2.4%	167	Thailand, Hungary, Malaysia, Turkey, Brazil, Colombia

Source: 'Multichannel movement' study by dunnhumby
Index: Sales of Multichannel customers vs. sales of one channel customers

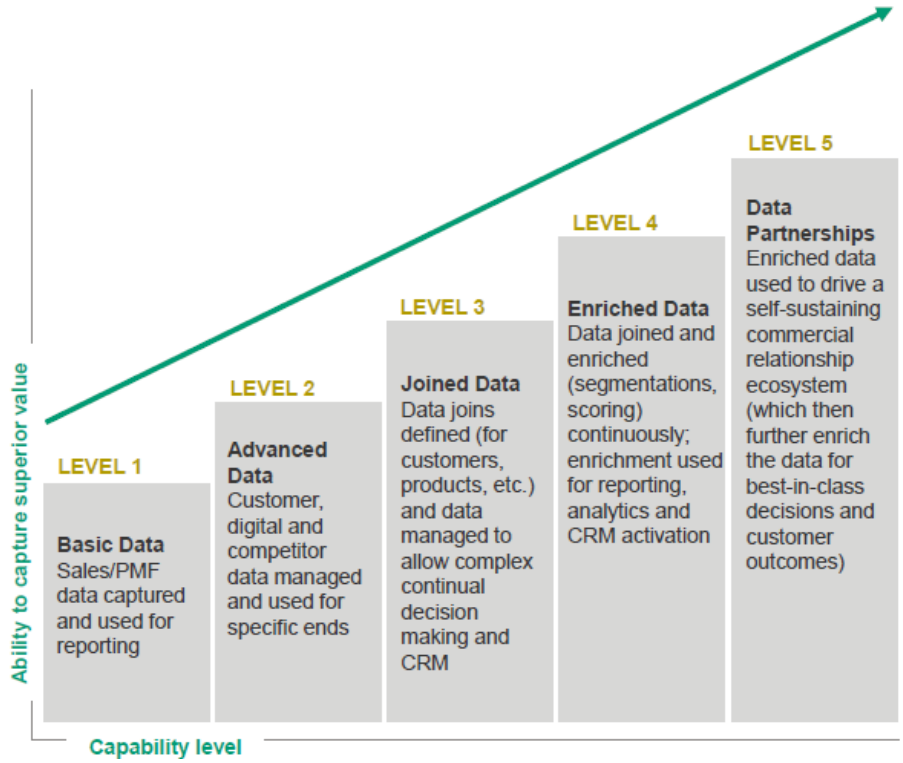
Thailand has a population of 68.1 million population, while 38 million are connected to the internet. Among those, there are 48 million social media and 83 million mobile subscribers. Hence, Thailand is one of the countries with the highest usage of smartphone. 85% said that 'Mobile devices are a central part of their everyday life'; 52% said that "They opt into push notifications for coupons and deals through mobile multiplatform of OSs and mobile application". To capture such a chain of change in customer behaviours, retailers and brands requires a shift in its Customer Engagement strategies.

With the aim of leading the change in Retailer Customer Engagement strategy, Tesco Lotus introduced its first mobile application in Thailand in 2015 along with its other digital connections platforms to customers. This gives Tesco Lotus the ability to engage customers, whether they are at home or on the move. Yet, there was still a need for in-store engagement and a technology platform that enables seamless engagement with customers.

Data source and methodology

dunnhumby has experience in working with big data for more than 25 years – well before the term was invented – and a wealth of expertise and experience in the design and implementation of big data strategies, as well as unparalleled and exclusive access to the data of 15 million club card holders and capture more than 650 million transaction per week in the biggest retailer in Thailand. The transformation to a data-led business doesn't happen overnight — it's a journey. It involves multiple steps to establish and integrate such a transactional, club card & product level, data with mobile / social media, location base information and offers to create a new building block of our new research methodology.

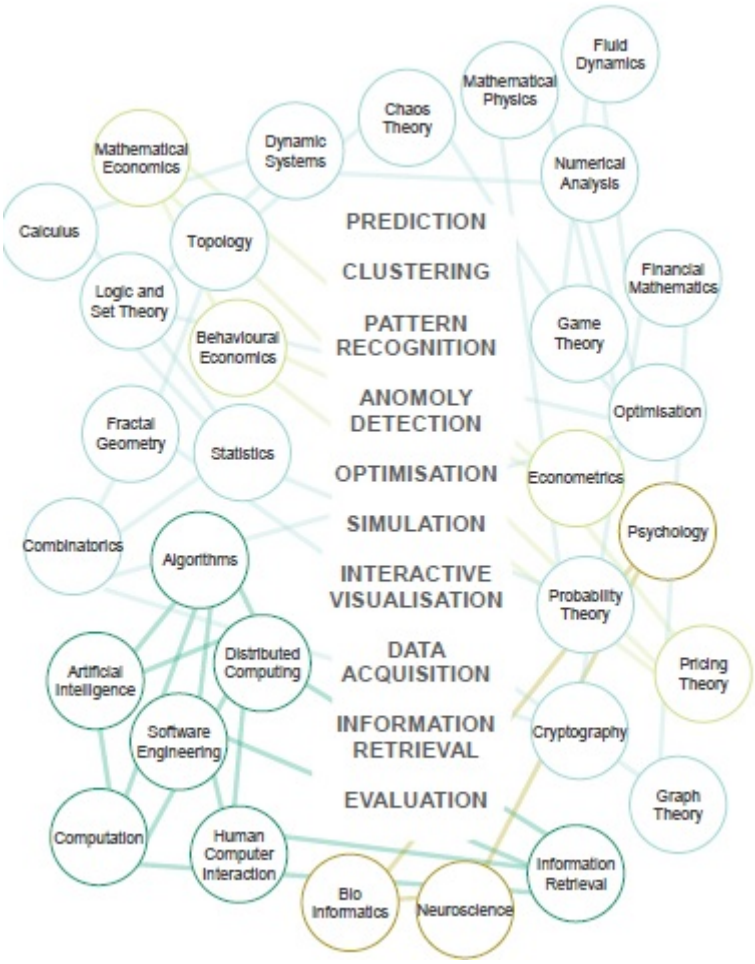
Figure 2. The data journey and its synergy to build new research methodology



These data sets are very large (big data) and complex, to use 100% of every single transaction (universe) requires speed & accuracy in coordination and that dictates the depth of methodology. Ultimately the range of insight is to answer the most important questions about customers, which enables delivery of insight-driven, personalised experiences, media measurement and effectiveness that keeps customers coming back, and generates an effective media planning and offer to brand.

The most important part of delivering such complex analysis is to use the fastest techniques to create full-scale models by developing computer programs that can teach themselves to recognize and predict patterns and relationships in customer data and generating the best data science.

Figure 3. The data science technique



Introduction to case study

With 15 million customer visit in stores every week, it is important to support shoppers to make their shopping easy across multiple platforms – offline and online; so called 'Omni channel' or 'multichannel'. It is also important to keep in mind how customers enjoy and interact with the mobile application in various platforms in Thailand. The main driving factors for such behaviour would also include how customers are being influenced and engaged through the social media platform.

In-store promotions still play a crucial role in driving sales in store; hence, communication to drive awareness is a key to success of the program. The standard communication platform has always been through retailer leaflet, TV, newspaper, and other social media. However, in recent years, TV and newspaper are no longer main channels that customers view for awareness of in-store promotions. Additionally social media has not played a successful role in engaging customers at the call-to-action stage.

Within the 15 million club card customers, there are 4 million customers who have downloaded the mobile application in order to shop online and get promotional information. The application also enables customers to locate stores through GPS, multimedia engagement directly from application or link to Youtube or VIMEO.

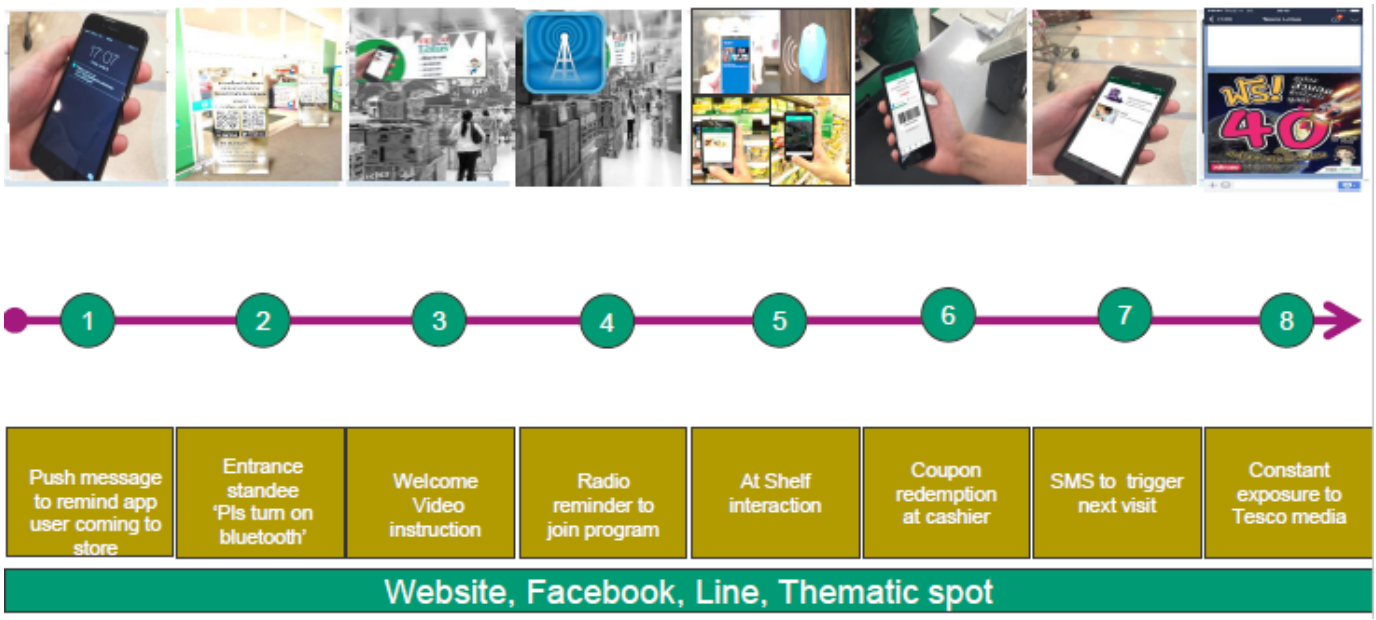
To complete the multichannel engagement platform, iBeacon was introduced. iBEACON is a technology that serves like an indoor GPS system and acts as a 'Content Delivery Platform,' serving interactive content such as video, reviews, or instant coupons the while customer browses through the store. This new technology platform enables us to talk to customers about the retailers' offer leaflets and, personalized promotions as well as drive cross purchase opportunity & improve footfall into low traffic aisles. It also provides an opportunity for brands to talk with their customer through VDO streaming, and social media.

iBeacon is based on Bluetooth low energy proximity sensing technology – it works by transmitting a universally unique identifier picked up by an application or operating system. The identifier and several bytes sent with it can be used to determine the device's physical location, track customers, or trigger a location-based action on the device such as a check-in on social media or a push notification, e.g. in-store promotion or trade offer.

iBeacon differs from other location-based technologies. iBeacon is only a one-way transmitter to the receiving smartphone or receiving device, and necessitates an application installed on the smartphone device to interact with the beacons. This ensures that only the installed application (not the iBeacon transmitter) can track users. iBeacons do not track the customers against their will, as they passively walk around the transmitters, only the app does the tracking. This data tracking with the link to club card customer ID can be utilised for campaign measurement; comprehensively synergizing such information with purchase data helps us in understanding customer behaviour better and how they interact between and within brands.

In delivering a successful campaign with iBeacon, it is important to engage customer through different media mechanics and communication throughout their shopping journey. This is to ensure that we educate the customers about downloading the app, turning on Bluetooth & how to benefit by using the app with iBeacon. Figure 4 depicts the media communication in-store and out of store and instructions on how to redeem coupons.

Figure 4. Example of customer journey and how to shopper redeem coupons



The iBeacon campaign has been introduced to customers in all stores; Hypermarket, Supermarket and Convenience formats in June – October 2016 with more than 5,000 iBeacon units installed across the country. The promotion offer is inclusive of both retailer trade promotions and brand promotion offers with the aim of serving different campaign objectives such as growing customer penetration, growing shopper basket spend, building cross purchase, and gaining more shopping frequency.

Figure 5. Example of coupon offer through iBeacon push notification

Measurement methodology

The campaign set up is to understand effectiveness from various customer perspectives. Those shoppers exposed to the campaign and those not exposed are identified. Then their pre-period behaviour is matched to identify test and control groups. The customer profiles are matched based on five main criteria:

- Shopping habit & Loyalty segmentation
- Price Sensitivity segmentation
- Preferred store region
- Customer spend and quantity at store level
- Customer spend and quantity at brand level

The basic measurement for the campaign is to understand how shoppers interact with push messages; Figure 6 provides the shopper flow for a clear view on opportunity and success. (See figure 6.)

In addition to push message and campaign interaction, it is important to understand change in customer behaviour based on the offers and the impact on the brands. To get into such a level of understanding, it is required to measure the campaign success on both Club card and non-Club card shoppers. The matched customers, (customer look-alike based on pre-period profile comparison) is being applied for test – control comparison..

Figure 6. iBeacon shopper flow

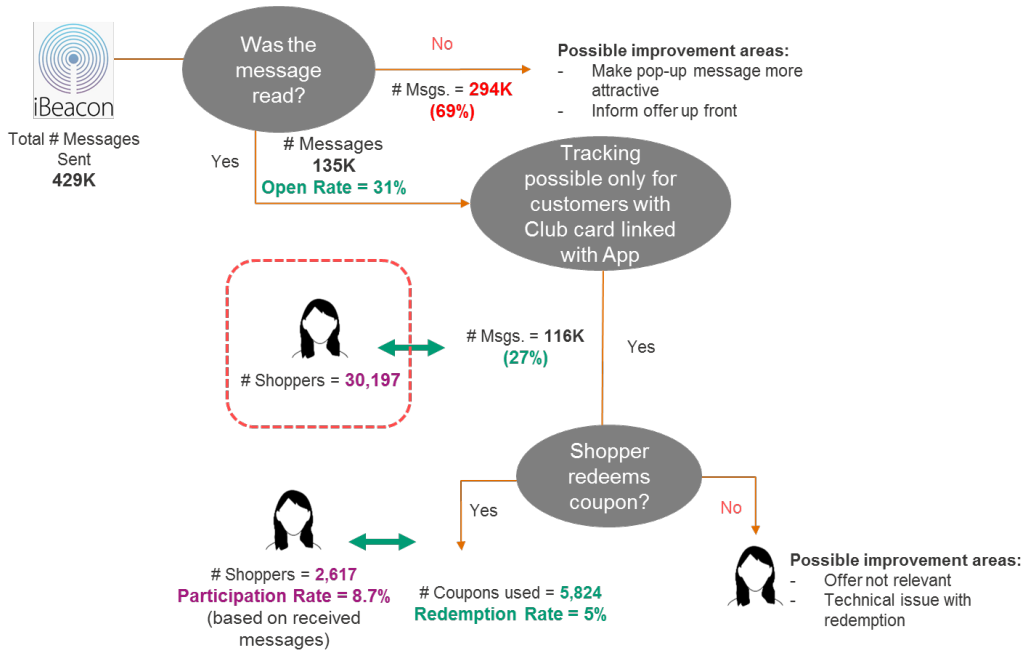
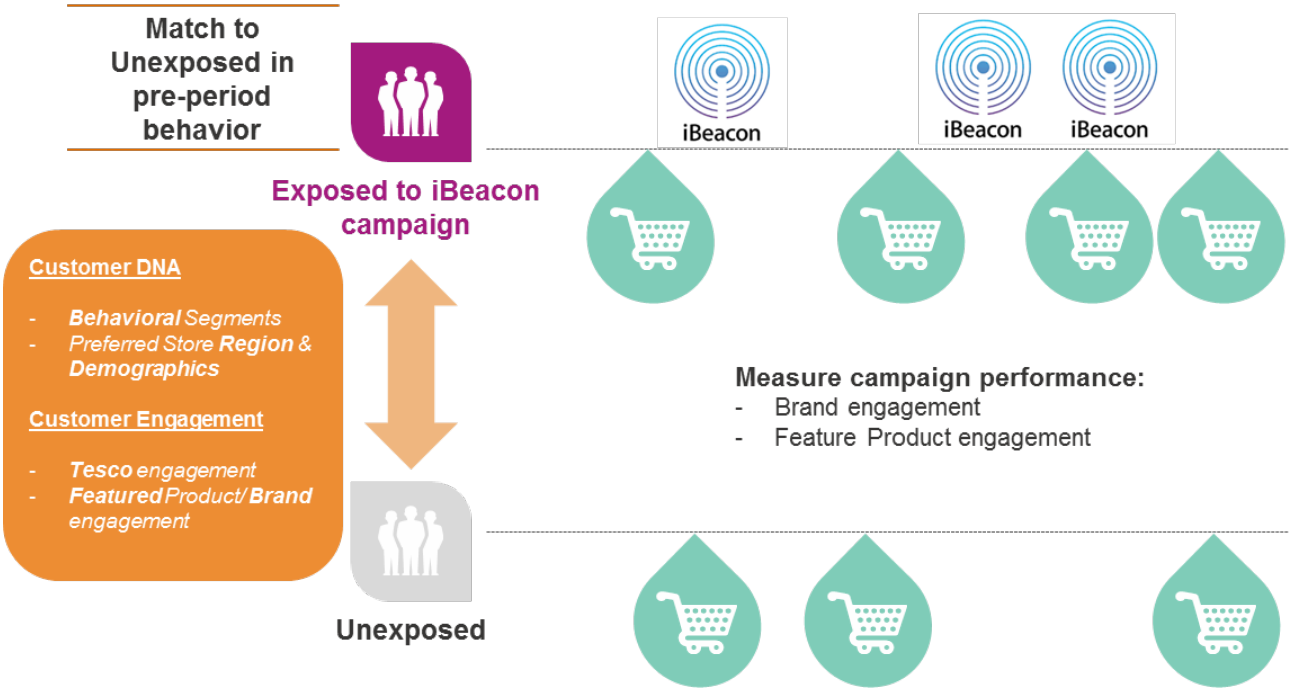


Figure 7. Measurement set up base on customer selection matching criteria



For longer term measurement, it is important to consider customers who were exposed to the media compared with customers who were not exposed. The resulting difference in behaviour tells us how the campaign has affected customers' behaviour.

Return on customer (ROC) is a campaign measurement perspective that means the long term return on investment delivered by the media campaign. In the post-recession era, it is very important for marketers to understand and justify marketing investment. Getting more customers to the brand is the sure fire way to improve business. These customers in the long run bring revenue to the brand which they would have not gained if they had not gained these additional customers during the media campaign.

To get the ROC for the campaign is a two-step process:

1. How many additional customers the brands gained during the campaign?
   We can answer this by comparing two groups of customers, those who were exposed to the campaign and those who weren't. We make sure that these two groups look exactly the same before the campaign (Test vs. control) so that any difference during the campaign period is because of the campaign exposure. The difference between conversion rates of buying the brand between these two groups can give the additional customers gained by the brand.

2. How much revenue will these customers bring to the brand in lifetime?
   This in turn can be answered in two steps.

   - How long will these customers stay with the brand?
     This we can estimate using the customer survival rates with the brand. Survival rates are calculated using the standard Kaplan Meier curve plotting. We take two years of customer transactions with the brand. We then find the first and last shopping dates with the brand for each customer. We censor customers whose last shopping date lies within three purchase cycle of the most recent date (cannot be flagged as having left the brand). We also censor customers whose last shopping date with the brand is also the last shopping date with the retailer. For everyone else, we plot the Kaplan Meier curve and get weekly estimates of survival rate with the brand.

   - How much revenue will these customers bring before they go away from the brand?
     Once we know on average how many customers stay with the brand for how long we can estimate their revenue. We estimate revenue from these customers quarter by quarter (if the purchase cycle is longer than the quarter we estimate it half yearly and so on). The formula for the revenue estimate can be written down as shown below:

$c = additional\ customers\ from\ the\ campaign$

$$b1 = average\ spend\ per\ customer\ in\ the\ brand\ in\ a\ quarter\ \left(\frac{annual\ spend}{4}\right)$$

$b2 = average\ spend\ per\ customer\ in\ the\ brand\ in\ 2\ quarters\ \left(\frac{annual\ spend}{2}\right)$

$b3 = average\ spend\ per\ customer\ in\ the\ brand\ in\ 3\ quarters\ \left(\frac{annual\ spend \times 3}{4}\right)$

$b4 = average\ spend\ per\ customer\ in\ the\ brand\ in\ a\ year$

$s1 = \%\ of\ customers\ still\ shopping\ with\ the\ brand\ after\ 1\ quarter$

$s2 = \%\ of\ customers\ still\ shopping\ with\ the\ brand\ after\ 2\ quarters$

$s3 = \%\ of\ customers\ still\ shopping\ with\ the\ brand\ after\ 3\ quarters$

$s4 = \%\ of\ customers\ still\ shopping\ with\ the\ brand\ after\ 1\ year$

$$Long\ Term\ Revenue = (Revenue\ from\ quarter\ 1) + (Revenue\ from\ quarter\ 2) + \cdots$$

$$till\ all\ customers\ stop\ buying\ from\ the\ brand$$

$$\begin{aligned} Long\ Term\ Revenue \\ &= (all\ additional\ customers \times avg.spend\ per\ customers\ in\ 1\ quarter) \\ &+ (remaining\ additional\ customers\ after\ 1\ quarter \\ &\times avg.spend\ per\ customers\ in\ 1\ quarter) + \cdots so\ on \end{aligned}$$

$$Long\ Term\ Revenue = (c \times b1) + (c \times s1 \times b1) + (c \times s2 \times b1) + (c \times s3 \times b1) + (c \times s4 \times b1) + (c \times s4 \times s1 \times b1) + (c \times s4 \times s2 \times b1) \ldots so\ on$$

We can see it is a summation of four geometric progressions till infinity. Using the standard formula for sum of geometric progression till infinity we simplify the above summation to the formula below:

$$\boldsymbol{Long\ Term\ Revenue\ (LTV) = c \times b1 \times \frac{(1 + s1 + s2 + s3)}{(1 - s4)}}$$

The above formula can be simplified further if we estimate using half yearly or yearly estimate instead of quarterly estimate.

If we take the above calculated LTV and use the media spend of the campaign we can easily get the return on investment or Return on Customers (ROC)

$$\boldsymbol{Return\ on\ Customers\ (ROC) = \frac{(LTV)}{(Media\ spend)}}$$

The result

dunnhumby, with its expertise in data science and analytics, measured the impact of the iBeacon campaign. The main result has shown a positive benefit on customer behaviour with the engagement of mobile applications and mobile offers. iBeacon has played a vital role to unlocking the potential of in-store communication such as recruiting new customers to the brand and category, generate sales uplift and drive cross purchase opportunity.

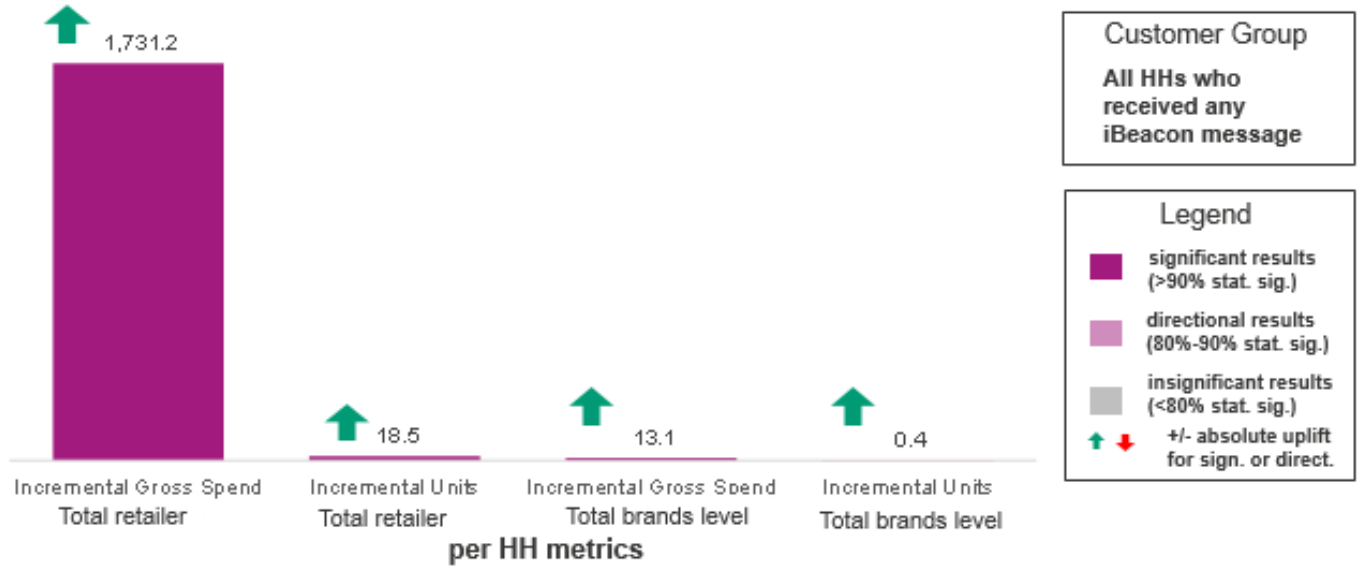
The preliminary result shows a participation of 8.7% (figure 8); of the customers who were exposed to iBeacon coupon offers. activated, 45% are new to brand, and they generate 17.6% basket spend uplift and 19.7% volume (units) uplift for the brands (figure 8-a).

Among exposed customers, there was also a 27.4% higher customer activation rate (figure 8-b) compared to customers who were not exposed. So, exposure to relevant offers based on location in-store works. In addition to pleasing customers, it translates into actual sales – an ROC of 1.4.

In addition, the campaign has shown stronger performance in Large Trading Format (Hypermarket and Supermarket) where there is a big opportunity to improve customer engagement through in-store communication. The campaign result has also shown clear evidence on iBeacon's role in growing customer spending in category and retailer.

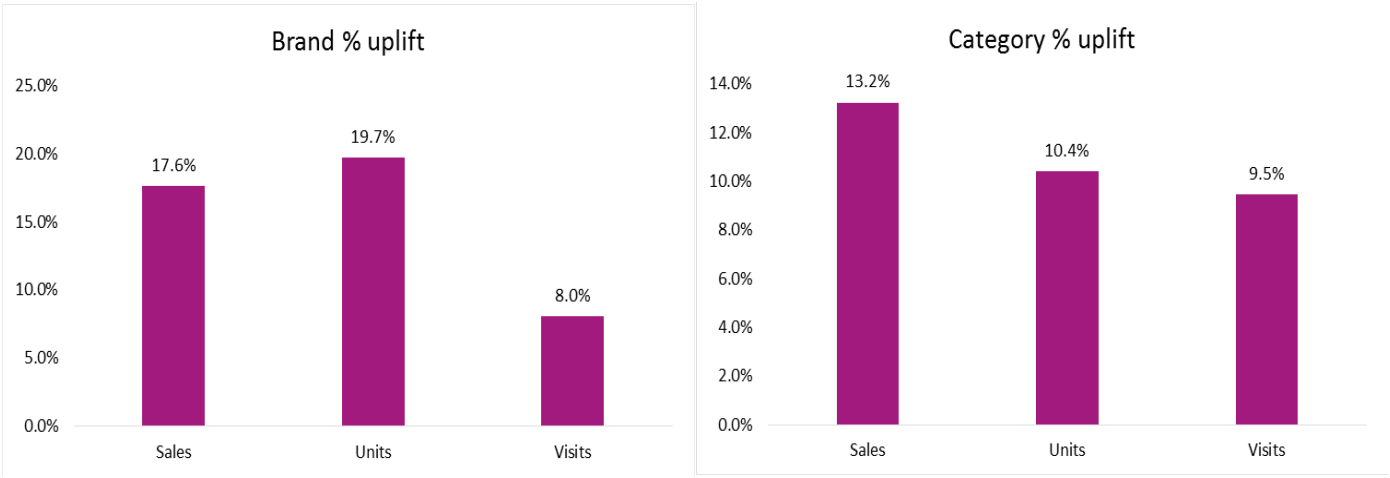
A sales uplift of 13.2% and a volume uplift of 10.4% was seen at the category level (figure 8-a). At the retailer level, the sales uplift was THB 1,731 and volume uplift was 18.5 units (figure 8).

Figure 8. Shoppers who engage with iBeacon campaign spent significantly more at retailer and brand level



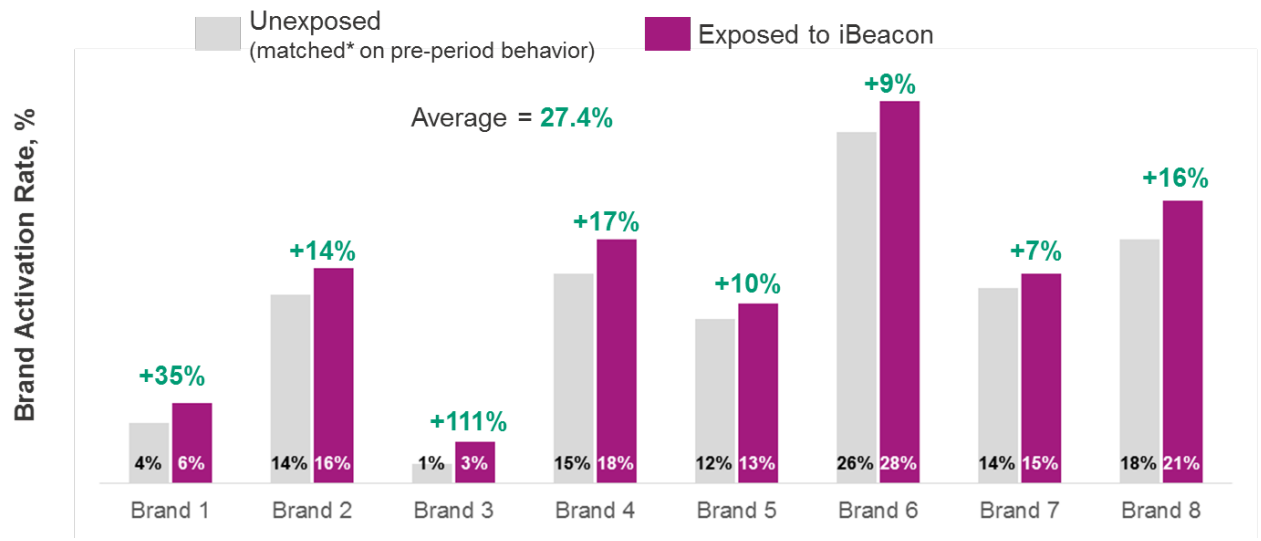
The figure shows the incremental sales/units generated per customer by the iBeacon campaign to the retailer and brands in THB/units over a 2 month period

Figure 8a. Shoppers who engage with iBeacon campaign spent significantly more at retailer and brand level



The figure shows the incremental sales/units generated per customer by the iBeacon campaign to the retailer and brands in THB/units over a 2 month period

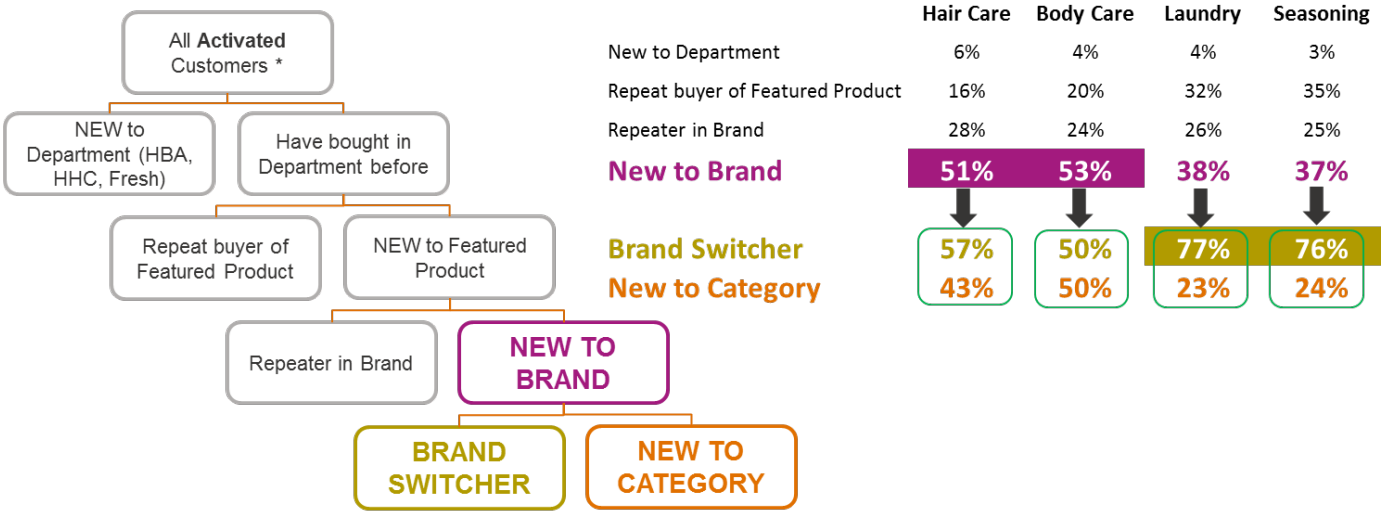
Figure 8b. Shoppers who are exposed to iBeacon show higher activation rates for the featured products



The figure shows the activation rates (activation is purchase of a product featured in iBeacon offers) for exposed and unexposed customers along with the percentage difference for the participating brands and the average across brands

It is clear that iBeacon offers play a different role in each of category depending on the level of customer involvement in a particular brand. Health and Beauty products saw the most success in recruiting new customers to the brand from outside the category; while Food and Laundry categories tend to gain more customers from brand switchers within the category.

Figure 9. Role of iBeacon offer to each category in order to recruit customer to category and brand

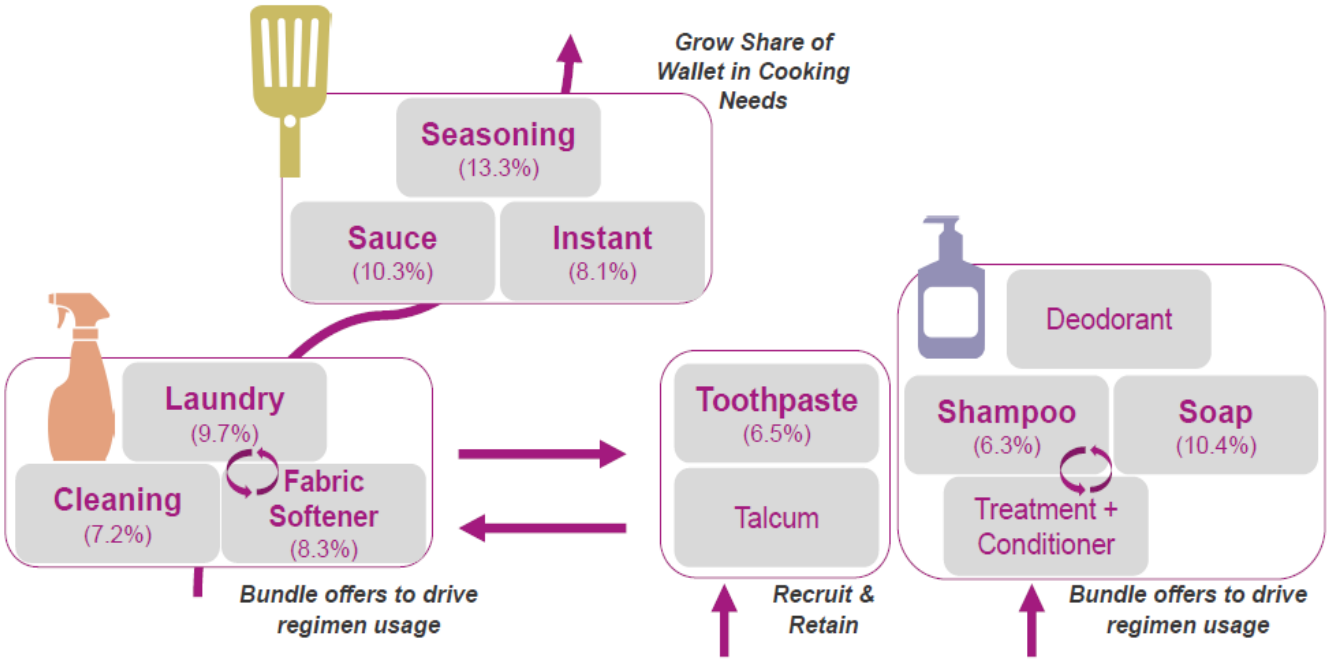


The figure shows the split of activated customers based on their pre-period behaviour. Customers who are new to brand, can be either Brand switchers (used to shop competitors' products in the category) or New to category (never purchased in the category before)

Conclusion

Purchase behaviour data is a real asset for better understanding of customers. Linking multiple data sets from purchase behaviour, Club card information and location based data could deliver personalised offers and shift customer behaviour by using mobile and innovative technologies. It is clearly transforming the way brands and retailers offer their benefits-of-product communication and trade promotion in order to grow basket spend (value and volume) and drive cross purchase opportunity (one more product).

Figure 10. Milestone to drive cross purchase opportunity between / within



The figure shows the opportunity for using iBeacon offers to drive cross-category purchase. The numbers indicate visit penetration of the category (Visit penetration = # baskets in category / total # baskets in the store)

The opportunity lies in scaling this up. By enhancing the customer experience and understanding how that can help brands grow their business, we are sure to improve the retailer's business as well. Customers need help in understanding how the iBeacon offers work – in-store activation helps them to achieve that. Breadth of offers and depth of discount help customers embrace the new geo-location based communication technology.

The new ROC KPI is used to validate investment efforts on customers in the era of change, which could be utilised and can lead the way to shift media measurement effectiveness.

The Authors

Teeradet Dumrongbhalasitr (Lek), dunnhumby, Thailand

Vijay Balaji M, dunnhumby, Thailand

PART 6
Discovering How to Design Brands and Great Customer Experiences for Everyone

Optimizing Customer Experience

Weiwei Zhang • Rebecca Chen

Introduction

This paper explains how to apply design thinking elements in the process of creating meaningful service experiences among the different customer touch points. How these different touch points are constantly changing in the China landscape, why these changes happen, and what brands should be doing in order to innovate and maintain a competitive advantage will be addressed.

The benefits of the process of discovery that design thinking research process provides and the key advantage of thinking by doing in China context is highlighted.

Key issues that are included include:

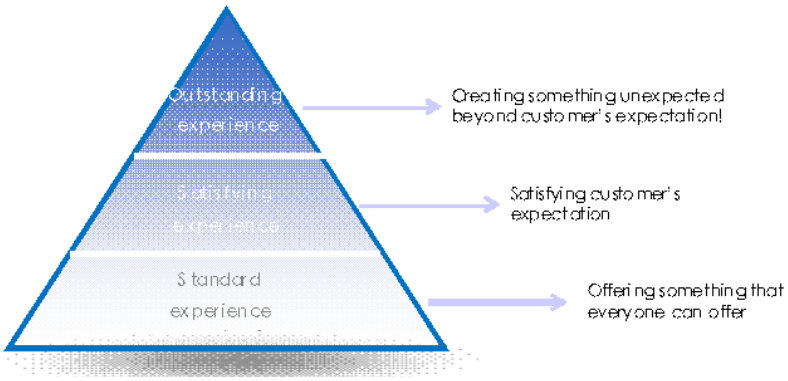
- Understand more what entitles a meaningful experience in the China market and what is driving change in consumer behavior in today's China market
- Learn about the models and skills of service design and how to successfully create innovative customer experience
- Learn about those typical scenarios and touch points where organization has the opportunity to make an impactful impression on a customer and develop the efficiency of sales.

The era of experience economy

As consumers have better tastes and higher demands, product competition is not just about the physical product itself, but also about all other aspects related to consumer experiences, such as packaging, service, brand concept and lifestyle.

It seems that experience has become something tradable instead of something accessorial for consumers, and may be sold for its own value as if it is an item for sale.

Figure 1. Let consumers encounter something unexpected in the era of experiences



For example:

Buy shoes, buy some coffee! A shoe brand store, focused on selling sports shoes, has dedicated a section of their store to providing customers with the opportunity to buy some delicious coffee while shopping. They call it playing with the senses in the retail experience.

Buy and experience beauty products as if in a bar! A beauty and cosmetics brand store created an in store experience as to make customers feel as if in a bar. Customers have access to cookies, tea, water, coffee,

and beauty products which showcased as if in the bar section. The more appealing the experience, the longer time is spent in store, the better brand perception.

Buy your prescription glasses calmly! A prescription glasses brand store designed the facility in such a way that the middle of the store had a long entertainment table for those individuals accompanying a friend/family member to decide on which glasses to buy.

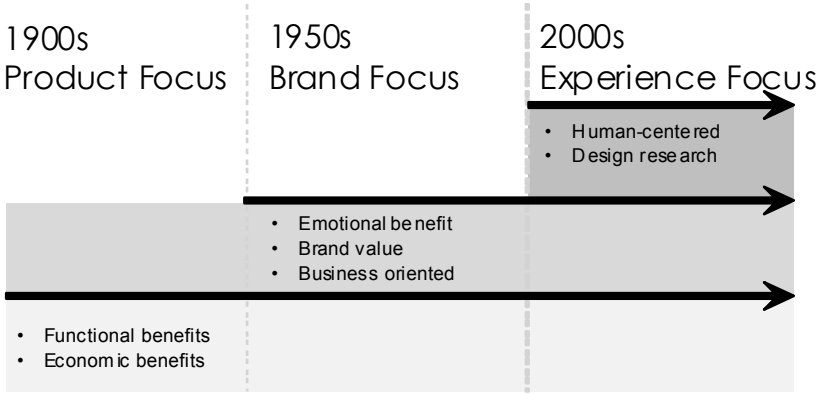
It has been proven that when shoppers are accompanied they may feel rushed to purchase, spending less time in the store.

Time to change the market research mind-set

In the past 50 years, researchers have already realized the importance of emotional benefits and lots of marketing decisions were made based on functional and emotional benefits.

Nowadays, with the arrival of experience economy, the consumers' need states have become unprecedentedly complicated.

Figure 2. Experience-oriented revolution



Trend 1: Comprehensive experience ≠ good experience

Offering a complete range of experiences in consideration of competition and various needs of the consumers may be a good idea, but it may also require investment of a huge amount of resources and lead to lots of wastes. Only by finding out the consumers' core demand and using your best efforts to satisfy such a demand could an unforgettable experience for consumers be created.

Companies should identify the position and core competitiveness of your experience: get the foresight and insight of consumers' core demand and expectation for experiences, and then offer the best experience to meet such a demand and expectation and build it into your core competitiveness!

Trend 2: Experience should be designed to fit with the living habits of people today

New media has changed the ways consumers access information and live. Thanks to the instantly interactive nature of the internet, there are many opportunities for improving client experiences. We should create interactions with our consumers, not only with electronic products but other products as well.

Trend 3: WOW experiences come from innovative methods and creative people

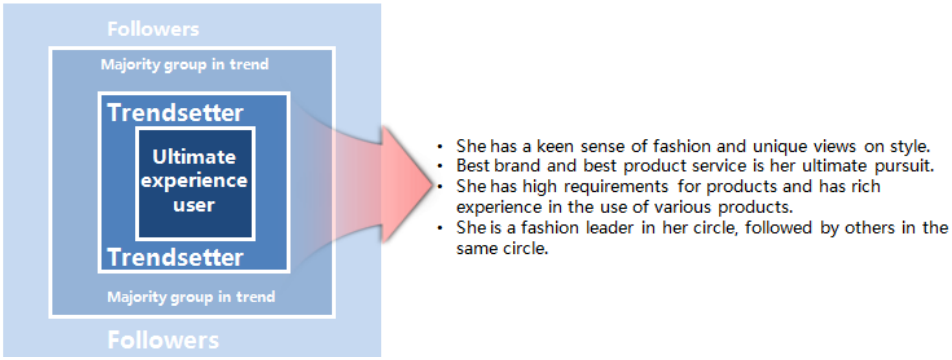
To find out what is beyond consumers' expectations, common consumer interviews are not enough. We need to use more innovative methodologies for inspiration and reference, such as ethnography, workshop or benchmark studies for the experience service industry.

- *Traditional methods*: We can only find out products and services that meet with customers' satisfaction through traditional need study or satisfaction study.
- *Now*: A new view is that we should explore some experiences that can make customers unexpectedly

touched. Innovative experiences are created by creative consumers and innovative methodologies (e.g. ethnography, experience study, workshop, etc.).

To find an innovative product that is beyond users' expectations, common consumer interviews alone is not enough. We need to use more innovative methodologies for inspiration and reference, such as studying the vanguard group and holding creative workshops.

Figure 3. Ultimate experience users



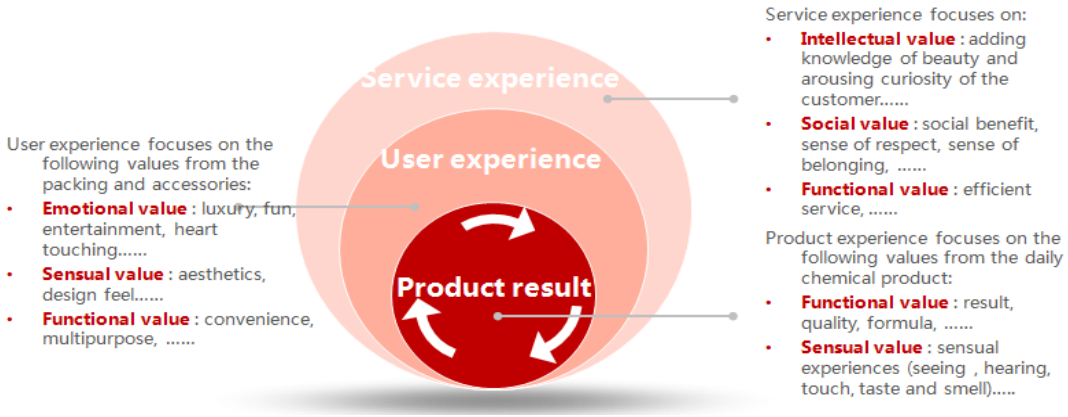
Trend 4: Assessment of the effect of an experience should be based on a real experience environment

Different from other products, experience is more about sensual and emotional value. Through simulation of a real experience environment, we can turn a "virtual" experience into a "real" experience, which may help to perform better assessment on the effect of the experience.

- *Traditional methods*: The method tested in the concept or explained in the service fails to reflect the real ideas of the experience and does not provide necessary interaction.
- *Now*: Make a prototype of the product for experience use and create a real environment with tools, materials or services to let consumers have certain time to personally feel the experience and actually get the real effect of the experience.

In conclusion, customer experience study is no longer focussed on a separated product or service process. It should be extended to all of the consumers' touch points. All touch points between the products and the customer are analyzed by the hierarchy thinking for seeking opportunities to elevate the value of the experience. (See figure 4.)

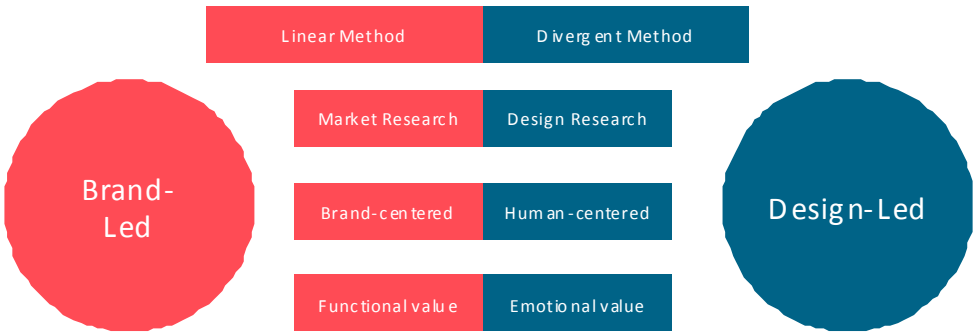
Figure 4. Build a system offering experiences to consumers through all touch points



Why design thinking?

Design thinking is a human-centered approach to innovation that draws from the designers' toolkit to integrate the need of people, the possibility of technology, and the requirement for business success.

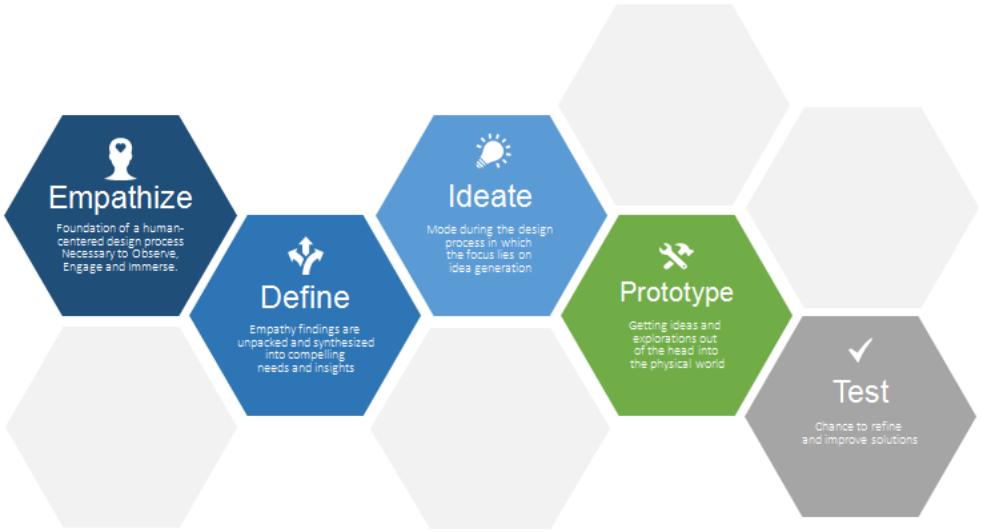
Figure 5. Brand-led vs. design-led



Design thinking in business uses the designer's sensibility and methods to match people's needs with what is technologically feasible and what a viable business strategy can convert into customer value and market opportunity.

There are five typical phases in the Design Thinking process.

Figure 6. Design thinking process

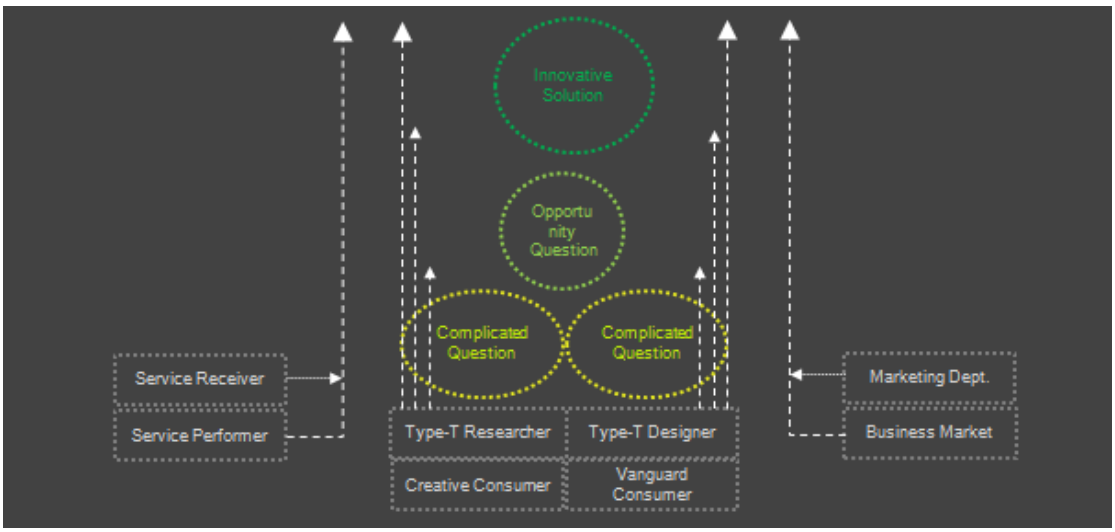


Empathize: Break down the process and action of experience and listen to the comments of the target group at different stages, in order to learn the key points in experience, so that many methods and many tools / ways can be used to strengthen the experience and feeling based on these key points.

Define: Identify all touch points throughout the whole process of experience can help to add / reduce / redefine / innovate different elements for experience. Based on problem list, we can summarize and analyze the touch points in experience to find the key points, identify the pain points and determines the right direction to go.

Ideate: Marketing researchers and internal personnel of the company work together through brainstorming and creative thinking to redesign the innovation process and get inspirations for a new direction and a new concept. In order to find out what is beyond consumers' expectation, we need to include the client's internal team, external type-T designer, creative consumers and vanguard consumers in our synergetic workgroup.

Figure 7. Creative workshop



Prototype: Make a prototype of the product for experience use and create a real environment with tools, materials or services to let consumers have certain time to personally feel the experience and actually get the real effect of the experience.

Test: Using quantitative or qualitative research methods to identify which prototypes would be most attractive to target consumers. Also, we can iteratively design and test the prototypes.

How to incorporate the design thinking process in traditional research processes

Based on the traditional qualitative and quantitative research, design research takes scenario, touch points, design and prototype into account. It would use variety of innovation tools, such as service blue print, problem list, service design, brainstorm and creative workshop, etc.

It helps researchers transfer the thinking mode from independent research to service design, which is more related to customer experience.

Figure 8. Market research vs. design research

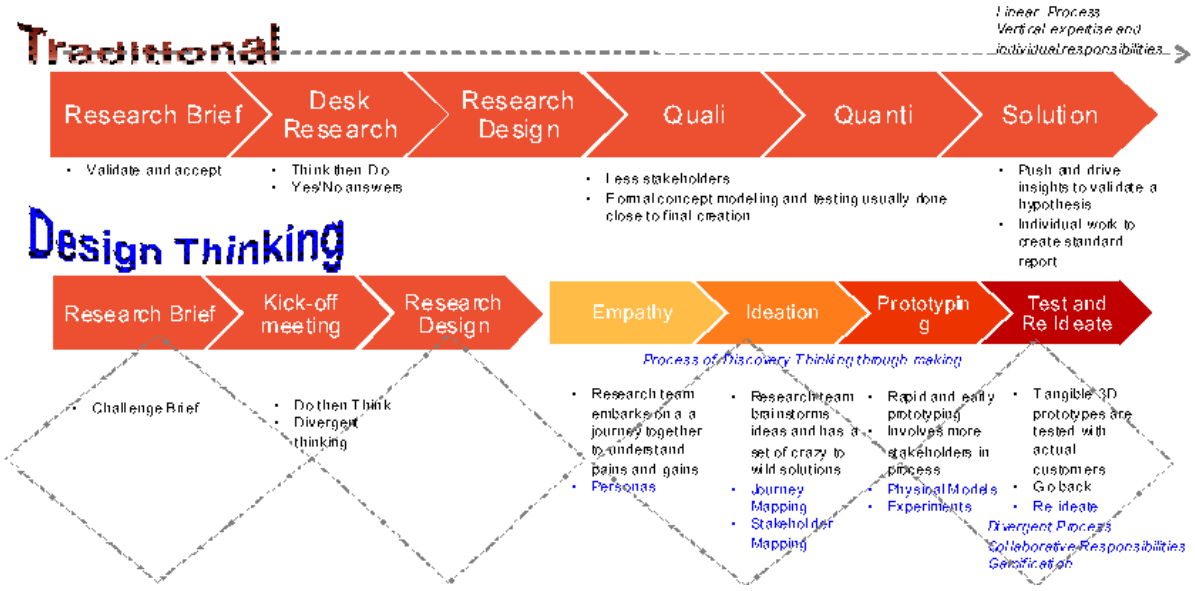
Market Research	Design Research
What people say they will buy	Why people said they would buy
Statistically valid large sample sizes	Small sample sizes
Focus What/who: Segments, trends and demographics.	Focus Why/how: Understand what target segment want to do with the product
Business insight.	Insight into the design, innovation and iteration process.
Identifying and uncovering target customer segments, preferences, opinions and requirements	Finding the needs people don´t even know they have
More quantitative. Tends to involve large numbers.	More qualitative. Observed behavior of users.
What people doing	Why people doing

Design research aims to perceive product features through user experience and directly guide the clients on how to upgrade their customer's experience.

Figure 9. Design research process

The traditional design disciplines focus on the designing of products' visual communication design, interior space design, product design, information design, architecture, planning, etc. However, the emerging design disciplines focus on designing for a purpose, such as design for experiencing, for emotion, for interacting, for sustainability, for serving, for transforming ...

Figure 10. From traditional research to design thinking research



Case: Beauty products consumer experience improving study

Research background

Market competition between mid- to high-end skin care products becomes more and more acute. Different brands are trying different ways to improve the immediate experience of their products, such as skin tester, value-added beauty service, guidance by beauty counsellor, beauty class, etc.

Beauty parlors are also changing and updating their experiences by providing more diversified products and devices to motivate customers to buy their skin care products. As there are more and more choices of brands, consumers usually feel numb and indifferent about those marketing approaches that are almost about the same kind. Consumers focus more on the product itself and expect to receive more professional and more surprising product experiences.

Therefore, we need to explore products, experiences and accessory tools that are effective in boosting purchase of skin care products.

Research methods

Based on the Design Thinking typical process, we used qualitative and quantitative research methods to solve the client's marketing problem step by step.

Table 1.

	1	2	3	4	5
	Empathize	**Define**	**Ideate**	**Prototype**	**Test**
Tool	• *Persona* • *Experience Map* • *Service blueprint* • *Purpose Value Chain*	• *Problem Matrix* • *Quadrant Analysis*	• *Creative workshop* • *Co-creation* • *Creative list* • *Concept list*	• *Co-creation*	• *Iteration Test* • *Prototype Simulation*
Output	• Insight Findings • Consumer Needs	• Problem list • Opportunity points	• Process design • Concepts	• Prototypes	• Assessment • Guide book
Methods	• Observation (N=30) • In-depth interview(N=30)	•	• Creative workshop (N=2)	•	• Ultimate user In-depth interview(N=6) • Quantitative pre-recruitment interview (N=72)

Case 1: Skin care service solution

Phase 1: Empathize

We invited over 30 consumers to experience at least 20 skin care brand stores. Based on the observation and empathy, we drew 22 service blueprints and some interesting trends of on spot experience.

Figure 11. Example of skin care service blueprints

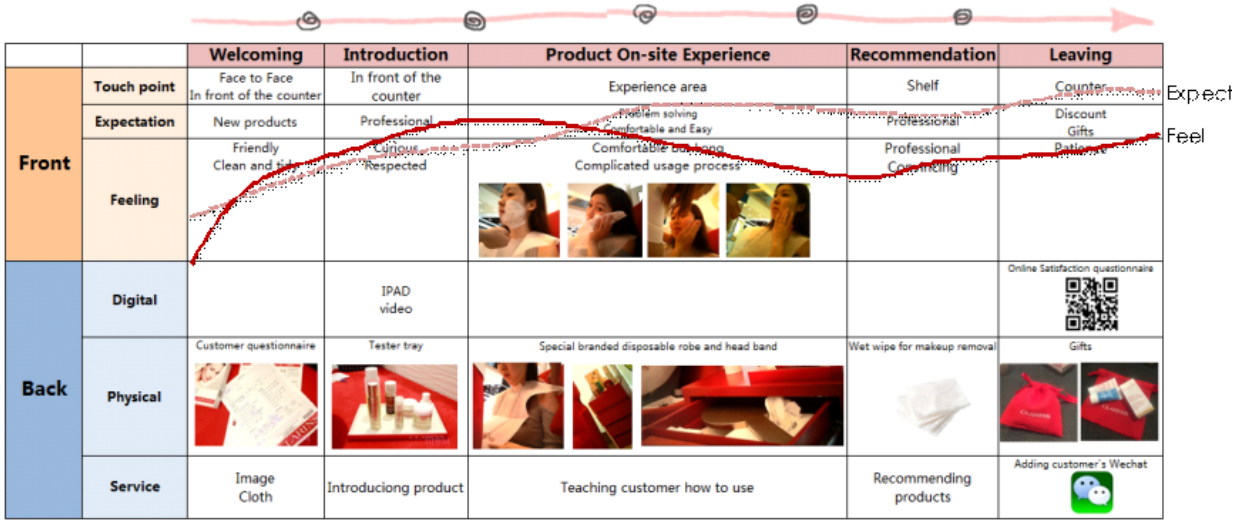


Figure 12. Skin care service blueprints

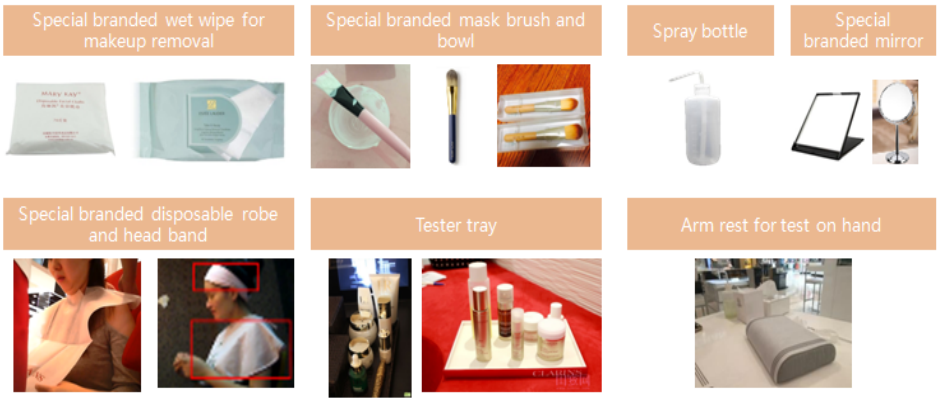
Phase 2: Define

Some interesting trends of on-the-spot experience

How can we utilize on-the-spot experience more effectively to touch the point of interest of the customers and maximize the sense of effect so as to promote purchase?

- *Trend 1: Diversified skin care service packages.* Making earlier appointments for the facial care service package at the brand store through brand's Wechat official platform.
- *Trend 2: Multiple formats of skin care solutions.* For example, more and more brand stores try to bring consumers with impressive and instant experience via multiple types of masks (paper mask or mud masks). Mask is a product which can better create enjoyable skin experiences than lotions, essence and cream.
- *Trend 3. Specialized auxiliary tools.* Brand stores try to build up professional images via offering different auxiliary tools during service process.

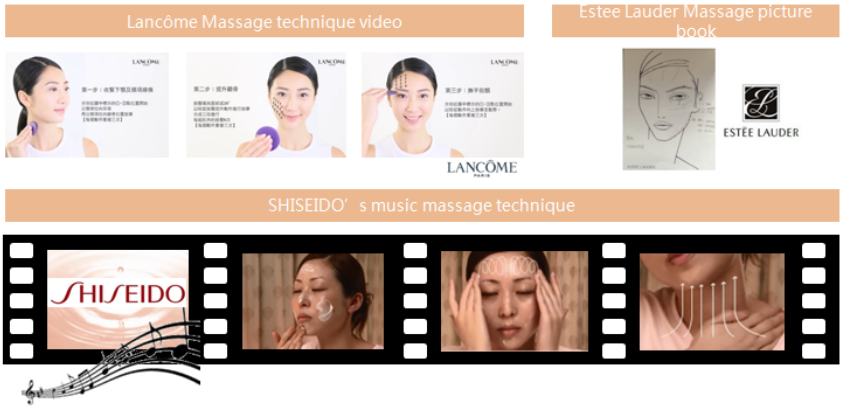
Figure 13. Different auxiliary tools



- *Trend 4: Hi-tech skin analysis*. More and more brand stores would use technical skin testing tools before skin care solution recommendations, for example, online questionnaire, skin tester, blue ray machine, etc.
- *Trend 5: Customized massage techniques*. Skin care sales people would emphasize massages so as

to enhance the consumer's product interaction so that the customer can feel the product result after each use and really enjoy it. For example, Estee Lauder's BA has a picture book in which the same facial picture is provided for the BA to draw the parts and directions of massage when explaining the massage technique to the customer on the spot.

Figure 14. Brand's massage technique



- *Trend 6: Self-serviced product experience*. Trial counter experience area (providing a basin and warm water for washing face, tissue for removing makeup, and facial mask brush) or encouraging customers to try on the product by themselves on the spot.

Figure 15. Self-service product experience

Phase 3: Ideate

Market researchers, client teams, designers, creative consumers, and marketing experts worked together through brainstorming and creative thinking to redesign the skin care service process and get inspirations for new directions and new concepts.

Figure 16.

Phase 4: Prototype

Make a prototype of the product for experience use and create a real environment with tools, materials or services to let consumers have enough time to personally feel the experience and get the real effect of the product experience.

The simulation environment included online invitation messages, standard service process, experience room, healthy drink, music, brochures, product, and other auxiliary tools. We aimed to create a real space where respondents would forget they were attending a market research interview.

Furthermore, we can interactively optimize the space according to attendants' feedback.

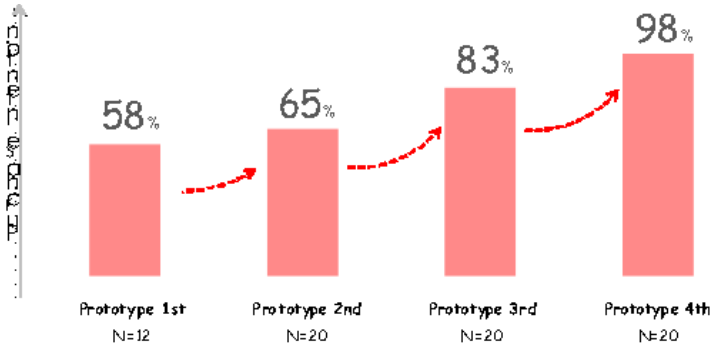
Figure 17. Service process simulation

Figure 18. Five sense environment simulation

Phase 5: Test

Using quantitative or qualitative research methods to identify which prototypes would be most attractive to target consumers, we can also iteratively design and test the prototypes.

Figure 19. Purchase intention level shows upward trend from the first to the third tests



Case 2: Apartments positioning study

With the crazy increasing price of real estate in China, more and more young people choose to rent apartments after they graduate from the universities. This research did not just tell the client how the young people choose and live in the apartments, but also guided clients on how to plan and design the apartment products for different types of target tenants.

Phase 1: Empathize

People from different backgrounds were invited to describe their living condition, life status, and some typical scenes from their life. First, we took photos from their house to get the feeling directly. Secondly, people were asked to write down the words or draw pictures to remind themselves of how they interacted with the places they live. Thirdly, they were invited to talk to each other about what kind of role the apartments played in their real life, and what the gap is between “Dreamed house” and “Real house”. By having completed all the steps above, we can better understand other people’s lives.

Figure 20. Description and feeling from consumers’ real life

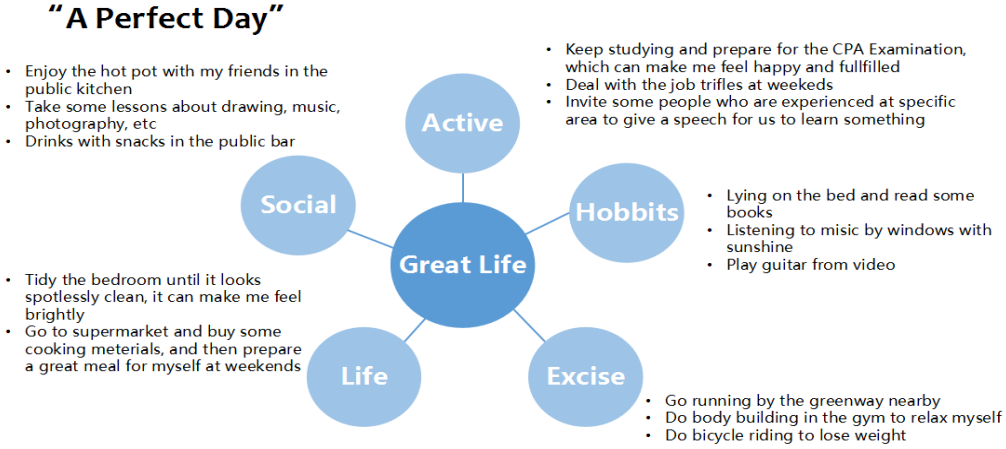
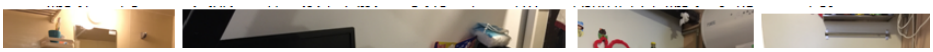


Figure 21. Observe the tenant’s apartment

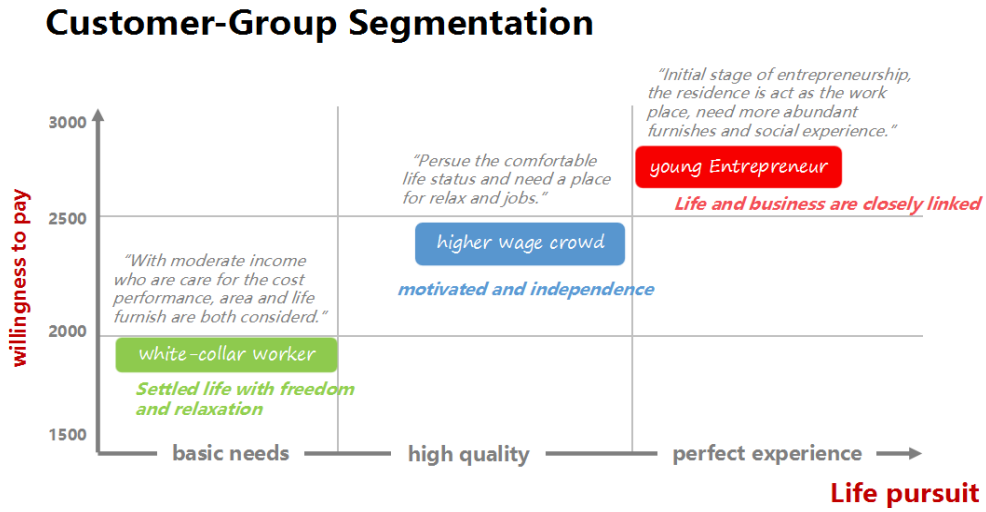


Phase 2: Define

According to the backgrounds and different needs from the residential apartments of consumers, we came to some definition to describe them precisely and effectively, which can help us to find out what kind of goods are suitable for them, even inspire them.

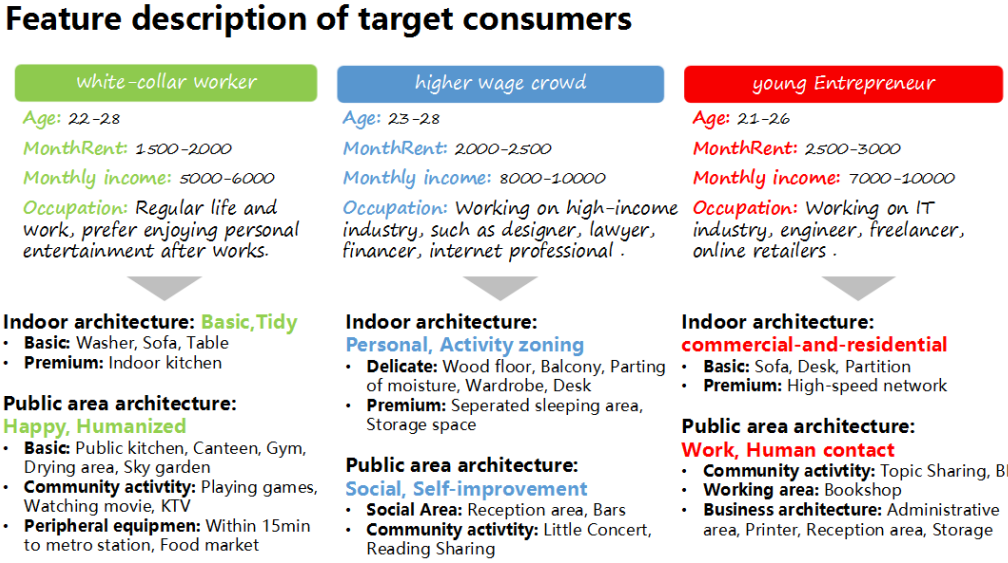
People's willingness is obviously related to their career experience. Then we classified them into three levels from the life pursuit. These two dimensions are positioned on the vertical and horizontal axis, and then we developed three labeling groups.

Figure 22. Customer-group segmentation of rental apartment



For each group, we conclude some keywords according to their description of ideal life — what they want and what they really need, and redefine the role of rental departments for them. We can see that there is wide variation among these three crowds.

Figure 23. Featured description of target consumers

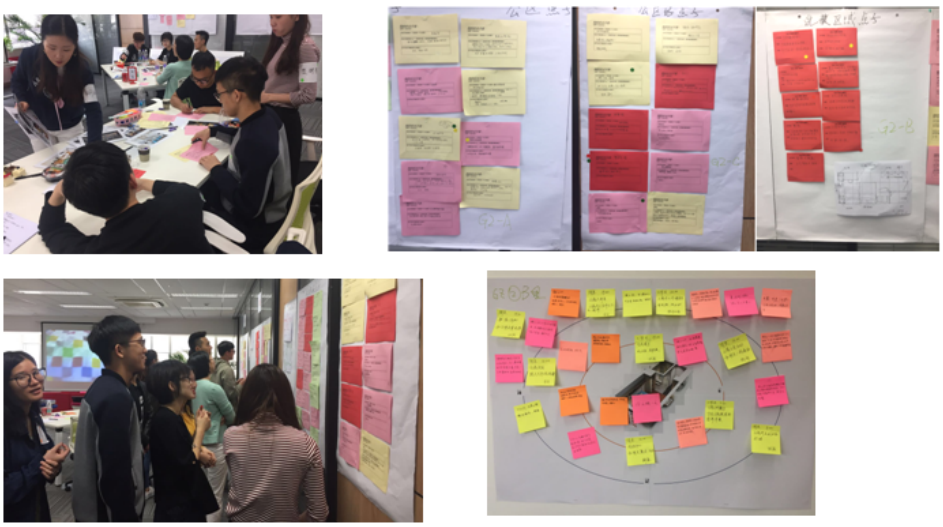


Phase 3: Ideate

Consumers are invited to work together and create some ideas about "Dreamed houses". They were inspired by colorful pictures and creative descriptions which were collected from some great apartment designs.

The most important principle of this process is that negative comments are not allowed nor is judging each other allowed. Without limitations and boundaries, participants are encouraged to inspire their divergent thinking and flourished imagination.

Figure 24. Creative ideas from workshops



Phase 4: Prototype

Use simple tools to create the prototype of their “Dreamed houses”. We offered colorful paint brushes, a lot of papers, scissors, glues, and any other raw materials for them to draw, to paint, to paste, and then created a house model which is visual and vivid. At this step, people could feel full sense of achievement and value, because they were doing something which may really change their life. And they also felt happy and interesting by creating something colorful and useful so quickly.

Restrictive factors should be considered by doing this, including manufacturing cost, payment capability, technical feasibility. So the professional designers and cost engineers are necessary.

Figure 25. Creating some prototype apartments

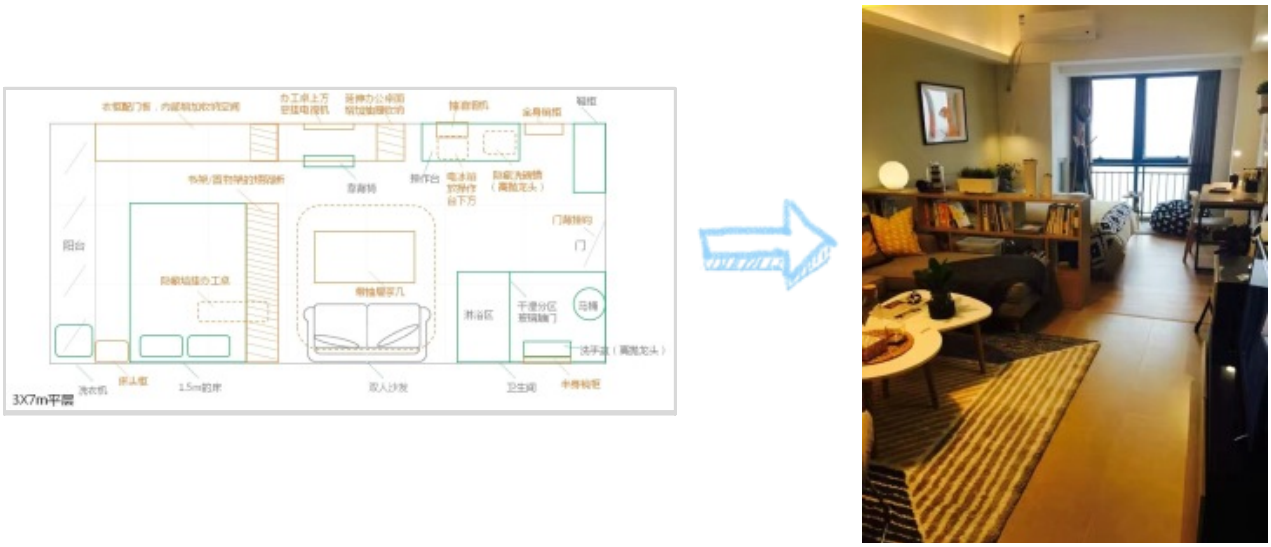


Figure 26. Testing different ways of partition

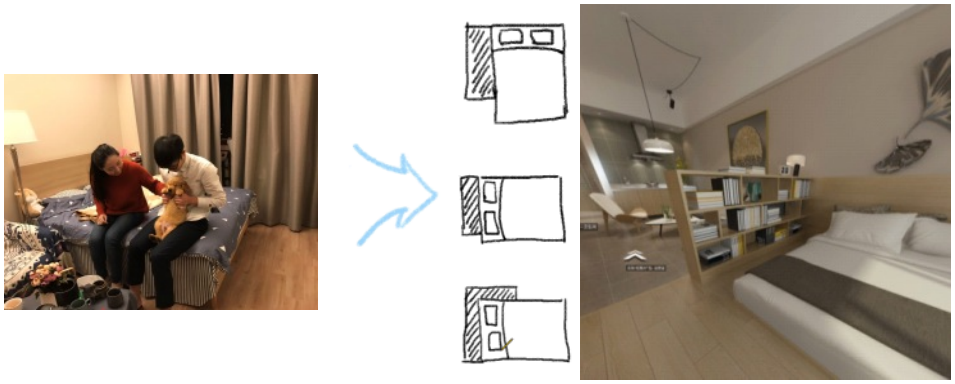


Figure 27. Optimizing storage experience

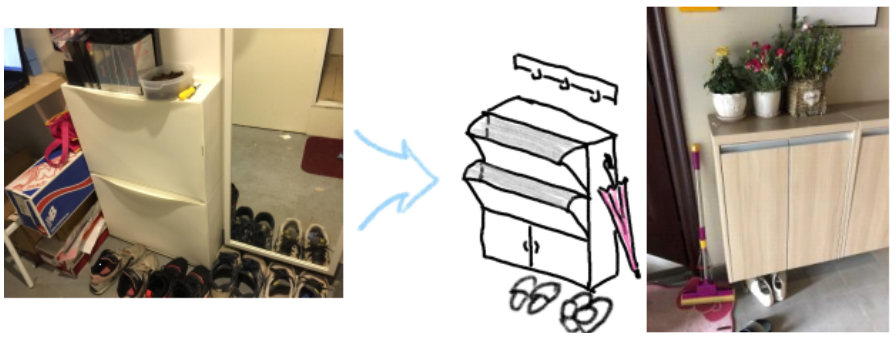
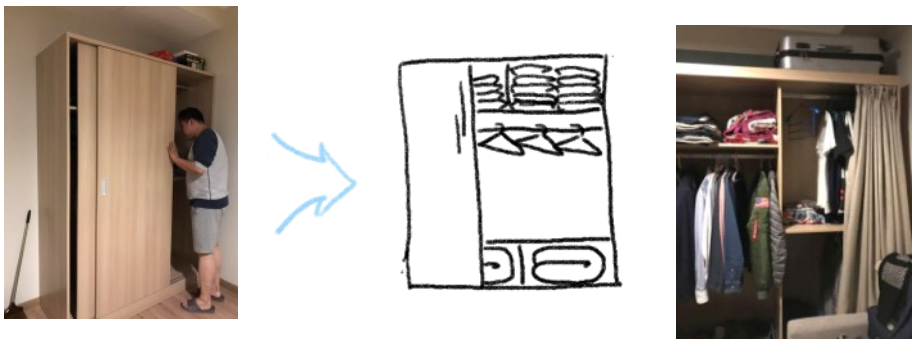


Figure 28. Creating better cooking experience

Figure 29. Ways to save costs: Replacing wooden closet door with door curtain

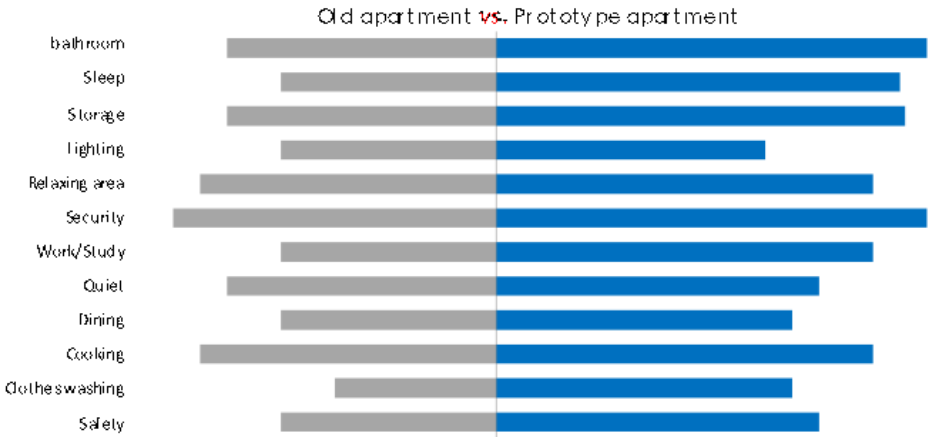


Phase 5: Test

Use a brief questionnaire applied with KANO model to do a small-sample survey. The aim is to identify which areas are basically needed and which are premium.

This quantitative method is rational enough for us to judge what is really needed for the target consumers and abandon some impractical demands.

Figure 30. Comparison between current apartment design and prototype



Conclusion

With the arrival of era of the experience economy, Chinese consumers have better tastes and higher demands on brands. What consumers want are not just the physical product itself, but also all other aspects related to consumer experiences, such as packaging, service, atmosphere, social value, etc.

Therefore, testing is no longer a simple product test or a concept test. Market research must transfer to Design Research. Design research is an integrated process which bundles the client and research team together tightly.

We strongly believe that design thinking will bring traditional market research methodology a new mode which can lead to better marketing innovations and solutions.

References

User Experience: The Ultimate Guide to Usability (Posted on March 8, 2013 by Matthew Magain in Analysis, Reviews)
Five Trends Impacting Customer Engagement & Support
Why Every ”Touch Point” Counts in Customer Service (2013-07 by Elaine, ElaineAllison.com)
How to Identify Your Customer Touch points? (Surveymonkey.com)
Design Thinking (Tim Brown)
Design as Art (Bruno Munari/Penguin Global/2009-05)
Creative Confidence (Tom & David Kelley/Crown Business/2013-10)
Change by Design (Tim Brown/Harper Business/2009-10)

The Authors

Weiwei Zhang, INI International, China

Rebecca Chen, INI International, China.

Gender Bias in Brands and Business

Using gender bias for good, not evil

Amy Fridlund • Nikki Feld

Introduction

Businesses and brands have an opportunity to grow market share by better meeting the need of the genders. But it can be challenging to see the opportunities when gender bias, particularly unconscious gender bias, obscures them. This paper pilots and presents an approach that quantitatively measures both conscious and unconscious gender bias across brand touchpoints as determined by consumers to reveal areas for improvement. We show that armed with those numbers and a strong dose of organisational awareness, brands and business have the potential to overcome a legacy of gender bias and outdo the competition.

The indications

Over the past 50 years the world has witnessed a remarkable rise in female paid employment and with it has come changes to gender roles and identity. A recent Pew research study of 30-44 year-olds showed that the more money women earn, the more spending they control. Sure enough, studies confirm that women today influence anywhere from 64%-85% of buying decisions and once male dominated categories are fast becoming female dominated ones and vice versa. Today females are accountable for an estimated $US12 trillion in spending. At the same time, research has found that the percentage of men who identify as the primary household grocery shopper has doubled in the past two decades in the US, as has the percent that are primarily responsible for child care.

What hasn't changed to the same degree is how brands cater to men and women. Men's and women's roles and cultural identity have changed, and they do not feel today that brands and businesses understand them. A 2004, 10-month multi-country study by Leo Burnett Worldwide, '*Miss Understood: She's not buying your ads*' resoundingly concluded that advertisers were failing to connect with women around the world. A 2009 book by the Boston Consulting Group, *Women Want More*, says, "…almost all women experience simple, chronic frustration with particular products and services (as well as some whole categories of goods), which seem to have been created and marketed by companies that have little to no understanding of who women today are and what they want." Over a decade later, despite attempts to rally industry, little seems to have changed. A recent study by the Futures Company on behalf of Unilever discovered that some 40% of women do not identify with the images of them portrayed in ads. This and other studies funded by Unilever found that 90% of females believe women are presented as sex symbols in advertising and 30% think they show a male view of women. In June 2016, in response to uncovering these discrepancies, Unilever initiated a business wide re-evaluation of how they depict women in their advertising and committed to presenting a more progressive vision of female identity in their communications, one that more accurately reflects the norms of contemporary society. It's a great start. That said, when we tried to find even just a couple of comparable statistics for men, we couldn't find any.

The inkling

Across touchpoints, including advertising, products, customer service and shopping brands continue to fail to truly understand and meet the often unique needs of each gender. In particular, we believe that the needs of women are underserved or often ignored altogether. In market research we continue to hear women lament that they feel ignored by financial services products which don't meet their needs. We hear women protest that automotive brands sales staff continue to patronise them in dealerships. And we hear women complain about feeling overwhelmed in consumer electronics stores. But the problem goes deeper than that. Even in categories where the end user is more likely to be male than female, we have to question why that is. We have to question the causality. Are fewer women using something because it isn't designed with them in mind? Or are fewer women using something because they genuinely aren't interested? This is part of a historical foundation of gender bias that is often unacknowledged and under-addressed.

The inspiration

We believe 2017 will be the Year of Gender. And while we think that it will be more about women, as suggested from what we've already discussed, we also think it will be about the changing roles and needs of men. We have seen the movement gaining momentum at the World Economic Forum at Davos with its gender parity initiatives, at Unilever with their #Unstereotype commitment to readdress how they depict women in advertising, at the Cannes Lions with the newly announced advertising judging criteria, at Capitol Hill with the Women's March on Washington protest. Although many businesses and governments are waking up to the gender issue, not everyone is 'on board' just yet. In fact, we suspect many businesses don't believe they have a gender issue in relation to their products and services and are unaware of the unconscious bias they likely harbour.

In May 2016, TNS conducted independent research on gender impacts in advertising which showed that when you accommodate for people at one extreme, results are also realised for those on the other end. In this case, we found that designing better advertising for females led to better advertising with men. We hypothesise that this is true not only for advertising, but also for other aspects of products and brands from product innovation to customer service to retail environments.

The idea

But brands aren't always aware of what they can do to better satisfy the needs of and connect with men and women. In fact, we hypothesize that brands unconsciously, and sometimes consciously, exhibit gender bias. What do we mean by that? Sometimes the way a product is advertised and / or sold, and / or designed and / or supported serves the needs and desires of one gender better than the other. We suspect that men and women perceive this bias more strongly in certain categories, with certain brands, and across certain touchpoints. We believe that the genders often have different needs, habits, usage and rituals which are better met by certain categories, brands, touchpoints – and worse by others. We suspect that brands create these experiences sometimes intentionally and other times unintentionally. We believe that the subsequent consumer frustrations and joys will be sometimes recognised as a gender issue, but oftentimes not. However, because business has been historically male-dominated, we expect to find more male gender bias in general – across categories, brands and touchpoints. Despite the powerful economic influence of women today, we expect to uncover categories where women are disproportionately influential but inadequately served: where they feel products are not designed for them, communication does not speak to them, customer service does not understand them and environments have ignored them. In this research, we hope to uncover, explore and better understand gender bias, and unconscious gender bias, in particular. We aren't sure if it exists, where it exists, to what extent it exists, or how much is conscious versus unconscious. So, we decided to find out more.

Unlike other studies of gender that examine the commercial sector of brands and customer experiences, our research will examine this topic through an economic lens and present the business case for better meeting the needs of each gender. A principle in human-centred design called 'designing to the edges' suggest that by better meeting the needs of people at one extreme, products and services can simultaneously better meet the needs of the other. We hypothesize that most brands create for the average – something that is acceptable to everyone but ideal for no one. 'Designing to the edges' rejects the notion of designing for the average, the lowest common denominator. Instead, it challenges brands and businesses to 'push the edge' and excel. We propose that by better meeting the needs of men and women, brands can unlock market share and increase profitability. Here, we propose an approach for uncovering this gender bias, identifying it across brand touchpoints, quantifying its impact, and highlighting the potential financial gains to be made by addressing it, by better designing to the edges.

The imperative

Consumers may not realise that their lack of connection with a brand may be the result of an ingrained gender bias, but they may feel its effect nevertheless. When a brand connects with consumers in this suboptimal way, it is a lost opportunity. While it may be a moral prerogative for brands today to challenge gender stereotypes, it is a business prerogative that they reconsider how they meet the needs of men and

women and win market share. It seems efforts to address the authentic differences between men and women have stalled, possibly in the name of political correctness or simply resulting from a lack of direction. So, although we may understand that equality does not mean 'sameness' in this post-feminist era, we still have a lot to learn about what makes men and women happy when it comes to products, communications, shopping experiences, customer service and employment.

As a result, we believe brands that take a step back, check their own biases, re-evaluate touchpoints in light of new understanding, and then re-design their businesses 'to the edges', are likely to unlock a large proportion of the female dollar and deliver to a greater part of the male wallet while they are at it.

The investigation

It is our intention that this research be replicated in other markets around the world. However, we have run the initial pilot out of Australia, a country with strong gender parity, as a stress test of the sensitivity of our approach. Detecting gender bias in Australia suggests our approach will be sensitive enough to use across the globe. Our research methodology is comprised of three phases: qualitative expert interviews, qualitative online bulletin boards and a quantitative survey. Currently, we have completed the first phase of the research, but we intend to present results and analysis of all three phases at the ESOMAR conference.

In Australia, we will examine two categories: quick service restaurants and mobile phone brands. We selected these categories for their equal purchase and usage rates amongst men and women. Note: in our study, we will be researching all aspects with both men and women, in order to obtain comparative measures with real meaning.

Phase 1.1: Expert interviews

In Phase 1.1 we conducted semi-structured qualitative expert interviews with gender and marketing academics, creative agency practitioners and 'c-suite' business executives from Australia and overseas, with the aim of understanding gender issues from both global and local perspectives as well as theoretical and practical ones. These interviews have served to validate our grounding hypothesis − that gender bias exists but that it is often unconscious. We have used these interviews to build our anecdotal understanding of how gender bias manifests today, develop additional hypotheses, and further our focus of exploration in subsequent phases.

The inquiry

While we have yet to complete Phase 1.2: Qualitative Online Bulletin Board and Phase 2.1: Quantitative Survey, our analysis will examine both the qualitative and the quantitative results by gender. Not only will we understand which brands are preferred by which gender, but by looking at the differences between the equity ratings of men and women of the various brands, we can quantify the impact of gender on brand preference. Any market share gap between males and females can be interpreted as an indication of gender bias. When its magnitude is quantified in dollar terms, potential profit growth to be gained by better meeting the needs of each of the genders becomes very clear. Furthermore, differences between ratings by control and exposure groups will help us understand the extent of unconscious bias versus conscious bias in the category, brand or touchpoint. We will analyse the touchpoint ratings for the various brands by gender and hypothesize as to potential causes for any gender discrepancies by leveraging learnings from the qualitative phases. Detailed analysis of the brand 'equity signatures' by gender, looking at differences in the size of 'opportunity' and 'risk' segments. will help identify strategic brand growth prospects and risk exposures.

Our hypotheses:

1. Brands unconsciously, and sometimes consciously, exhibit gender bias.
2. Brands create these experiences sometimes intentionally and other times unintentionally.
3. Brands aren't always aware of what they can do to better satisfy the needs of men and women.
4. Consumers are often unaware of gender bias; it is unconscious.
5. Consumer frustrations and joys will often not be recognised as a gender issue.
6. Men and women will differ in their ratings of certain categories, brands and touchpoints.
7. These differences in ratings derive from gender differences in perceptions, habits, usage, needs and rituals.
8. Men will display higher overall ratings across categories, brands and touchpoints. Conversely, women will display lower overall ratings across categories, brands and touchpoints.
9. We hypothesize that many brands create (advertising, products, services, environments) for the average – something that is acceptable to everyone but ideal for no one.

Through this research, we will gain a better understanding of how well or poorly categories, brands and touchpoints connect with each gender, and where growth potential exists. Importantly, our approach will enable us to gain a previously unseen perspective on the issue by quantifying unconscious bias. Uncovering this contributor to the status quo will give us a distinctive vantage point and empower us with a unique ability to help brands address brand touchpoint-related gender issues with clarity and confidence. Christina Habib, Vice-President for Refreshment at Unilever acknowledged a similar effect regarding their innovative approach to addressing their advertising, when she said:

"I've seen a lot of work about communicating with women in the past, but when we use the same approaches we come up with the same answers. This is the first time I've seen insights about women that reflect an understanding of the complexity of cultural factors that influence female identity, the cultural trajectory and dynamics at play. It tells us what we as a company of great brands can do to stay relevant and participate in the conversation. In acting on this new consumer understanding, it's our ambition to empower our consumers and employees to define the new social and cultural norms by which they live. This piece of work not only allows us to see women differently, it is acting as a clear call to and path for action."

The indications

As of this writing, we have completed phases 1.1 and 1.2. In phase 1.1, we investigated how marketing executives think about this issue and when, where and how businesses target and address the genders. Let us share what we have discovered so far.

Analysis of our in-depth interviews reveals that the gender issue in relation to brands and businesses is polarised in Australia. Some businesses exhibit a confident, proactive approach to managing their brands in relation to gender, and others appear to be ignoring gender altogether. We have summarised their attitudes below and allocated each attitude a descriptor for easy reference throughout the rest of the paper.

Business beliefs and attitudes that are barriers to change

Brands that ignore the gender seem to do so for a few of reasons:

1. They are blind to the issue and content with the status quo: *Oblivious*
2. They do not recognise continued gender bias. They feel the issue is a 'thing of the past' or that it is being adequately addressed: *Ignorant*
3. They are fearful of additional work: *Apprehensive*

Of those that do acknowledge it, the brands that don't act, do so for several of reasons:

1. They don't think it is a 'big deal': *Minimising*
2. They believe that there are only risks to addressing the issue and no potential gains. *Sceptical*
3. They believe that the risks outweigh the gains: *Risk Averse*
4. They believe that brands only in-authentically advocate for women to self-serving ends. Furthermore they believe that the only option open to brands is 'femvertising' — making an overt statement and communicating a controversial position: *Cynical*
5. They don't understand it- where and how it manifests: *Lost*

Of those that do feel they understand it, but don't act, do so for one reason:

1. They aren't sure why it is exists and therefore are not sure what to do about it: *Curious* (They may not appreciate the differences between men and women.)

As you can see, varying conditions create a framework of understanding, and a number of attitudinal typologies into which businesses can fall. (We expect consumers will organise into a similar framework.) It is likely that some businesses work through the typologies in stages as a stepwise progression, but we do not see this as a prerequisite to change. Let's discuss these typologies in more detail.

Fundamental to the framework is an acknowledgement of difference between the genders. Here, we build on work by Jane Cunningham and Philippa Roberts of consultancy PrettyLittleHeads, who have profiled differences between men and women and highlighted their importance for businesses in a structure that reconciles them. We believe the lack of acknowledgement of the differences between men and women, and thus the invisibility of the issue from both the business (and likely the consumer perspective) has a strong influence on its perpetuation. The classic example from the automotive category of women complaining about uncomfortable car seat belts underscores this issue. For a long time neither business (nor consumers) realised that this was a gender issue. Not acknowledging the differences between males and females, whether physical, cognitive, psychological or otherwise, makes it difficult to see where gender bias still exists in our daily lives. This myopia in both our professional and personal lives is the first barrier to change.

But businesses that fall into the first typology − 'Oblivious' − do not even look at the issue; they are blind. They fail to even realise that it is a potential topic to be addressed and the momentum of daily business propels them past the issue. Here, we find inertia perpetuating the norms of the past. Marketers in local offices are often presented with global advertising mandates with broad marketing targets, from far flung overseas head offices, which end up playing to stereotypes and old tropes. As one executive said, "There's no 'pause', there's no opportunity even to examine the topic."

Other businesses, even upon reflection, seem to be challenged to see gender bias in their businesses and need help understanding where it exists and how it manifests. These are the 'Ignorant' businesses. We suspect a large number of businesses fall into this category. But recognising gender bias is not an easy task, and brands have few methods or tools to help them examine themselves in order to overcome this hurdle of 'knowing the unknowns'.

Based on our discussions, the 'Apprehensive' typology seems to be found more often in larger companies. Executives in these organisations seem already challenged to consistently develop campaigns that deliver to expectations and reluctant to add additional gender criteria- another layer of complexity. It's not surprising that gender issues are put in the 'too hard basket'. When we do see such organisations challenging advertising gender norms, it appears to be as a result of directives from upper management.

Multinational companies seem to exemplify the 'Sceptical', 'Risk Averse' and 'Cynical' typologies. With them, stringent brand management accountability can make executives overly cautious, leading to few changes to targeting and advertising creative that speak to gender in a way that is in keeping with the times. Executives we interviewed said that these organisations are often scared to more directly address

the genders (typically the female audience), especially when it came to advertising and communications, for fear of public backlash (like the recent negative PR received by Audi for its Super Bowl LI ad) or alienating their current audience of buyers.

"I think people are scared to get it wrong, so I think it's easier to stay at the façade-level, because…if you get it wrong you'll be shut down for it. So, certainly there's a fear there of if we do take a bit of a risk and try to understand it and get it wrong, are we going to be seen as being contrived or trying to manipulate people, rather than genuinely understand them."

The issue has additional complexity:

"Brands don't want to be controversial or famous for the wrong reasons. They are worried about upsetting people, especially middle Australia. It is very hard at the moment to understand what people really value. Because we can talk about equality, and it's the right thing to do. And it's very politically correct… But if you don't agree with it, you don't talk about it – but you still secretly make decisions based on your beliefs. It's 'wrong' to say you don't believe in gender equality or gay rights, so no one will say it. So, it's a silent issue. But brands are worried that there is a 'silent majority' of people who value and want to see traditional family formats. So, putting two mums in a commercial is, maybe, an unnecessary risk for the brand, as opposed to just showing a [typical] family."

What more, these businesses rarely saw any commercial advantage to be gained by better addressing the genders. They recognised benefits only to be in the form of goodwill, and being seen morally to be 'doing the right thing.' Some businesses also seem to believe that their only option to addressing the gender issue in advertising is to make controversial or overt gender statements — 'femvertising' — thus exposing their businesses, both internally and externally, to enormous risk. They fail to see that speaking to the genders in authentic and subtle ways is an option. Not surprisingly, these attitudes and perceptions present strong disincentives to business and are significant barriers to change.

All of this is not to say that global brands are more apt to be 'behind the curve' on the gender issue. At the other end of the spectrum, dealing with the issue in very sophisticated ways, we also find the larger global brands represented. One creative director we spoke to said that his automotive client is actually one of the more gender progressive of his clients. He believes that the global financial crisis has helped to eliminate a lot of gender bias, amongst other inequities, having induced marketing with greater financial accountability. As a result, he finds that targeting and channel execution decision making is now very data-driven. Although he points out that gender is just one demographic criterion amongst many, as well as a strong dose of psychographics that, for example, contribute to how creative is tweaked in channel executions. Alluding to how progressive his automotive client is and playing on the cliché, he said that if his client found that women were more interested in engines and men were more interested in colours, then his client brand would market to them that way.

In the 'Lost' typology, we find organisations that are wedded to staid gender depictions that lag behind the times. As one MD commented, "I think we're working at [superficial] level. We're not even close to digging in and understanding what the true motivators are. For women, I think 'what are the genuine insights' and 'what are the motivators that are going to drive a behaviour change', rather than to the assumptions. And that's based on deeper insight, but that's rarely done well…[We need to] stop with the pre-assumptions on gender." Perhaps, pre-#Unstereotype Unilever might have fallen under this typology, having at some point acknowledged that gender bias existed in their business, but realising that they didn't know exactly what they were doing to foster it. This, perhaps, lead to the intensive research they conducted. As one MD of a creative agency said, "I've worked on campaigns for beer companies that have tried to launch a women's beer and it's probably failed because it worked on clichés as opposed to true understanding." While these businesses acknowledge the gender issue, many don't understand how or where it manifests or why – and need help.

Finally, there are those organisations that recognise gender bias — can see where and how it exists in their business — but don't know why. These 'Curious' business are uneducated on the differences between the genders. They might not understand that men are more apt to exhibit more logical, linear left-brained thinking versus females who are more whole-brained, or that men are more interested in control, whereas females are more interested in connecting. Many business executives found it difficult to articulate gender differences both because the topic can be so controversial, but also because they weren't clear on the commonly accepted gender differences.

Notably, organisations that *do* have a confident grasp of targeting and addressing the needs of the genders have a clear handle on these gender differences. One VP we spoke to said she is very aware of how women think and how that impacts on their creative strategy and advertising messaging, since women are the primary purchaser in the category. She notes that the competition still advertises with a male approach, connecting to only a male audience, and they can't understand why her brand does so well. Despite the fact that life stage is the biggest determinant in the category, with a 50/50 user base, it was clear that consistently finding better and sharper ways to meet the needs of the genders was a priority for her organisation. She went on to explain how all customer facing staff in her organisation receive 'customer care' training where, for example, men learn how to approach women, not to 'tower' over female customers and instead sit down next to them, etc. This organisation was also unafraid to look at their customer experience survey data split by gender and were willing to examine their business critically to understand where gender issues, amongst other systemic issues, might be arising. When it came to the product delivery environment her organisation was conscious of 'designing to the edges' from gender, age and ability / disabilities points of view. As such, it is clear that educating businesses and employees on gender differences is key to change. Perhaps, it should come as no surprise that a businesses' internal organisation along gender equality lines appeared to have a trickledown effect to the customer level, as well. Again, female executives noticed that their businesses were more apt to better address the needs of the genders as more females earned positions of leadership in their organisations.

However, the responsibility for not better addressing the needs of the genders can't be thrown entirely on brands and businesses. Industry creatives that we spoke to readily acknowledged that their own personal perspectives and biases were apt to creep into their work. A VP we interviewed said that one of her biggest frustrations was the lack of female leadership agency-side:

"I have been pitched to by a bunch of agencies and in my 6 years, here, I have seen only one female creative director- and she got fired. They're all male… And you can be the greatest creative director in the world, but you just can't get inside a woman's head. They just can't physically understand it. The frustrations I have agency side are far bigger than the frustrations, I have with my operations team about the guest experience. Because they're on the ground; they're sitting at the desk. They are so in touch with what is happening at the guest-level, if something doesn't work, they're going to tell you right away. Whereas these guys sitting in their ivory tower just have no understanding of a female experience."

Female CMOs, in particular, noted this and were frustrated by the distorted lens that they felt this brought to the marketing of their brands. A male creative director even noted counter-productive gender dynamics that can occur with an all-female creative team working under a male director.

But change is possible. Some of the creatives we spoke to said that, depending on where they join clients in the development process, they sometimes are in a position to challenge targeting and creative approaches. Motivated by the opportunity to tell new stories or present a fresh perspective, they can be a lever for change. We also found that using 'Design Thinking' also can be helpful. Because the process 'questions everything', it can be useful in challenging unintentional gender norms and revealing unconscious gender bias.

In the end, brands that are able to confidently navigate the gender conversation in business seem, in part, to be able to do so because they have moved beyond the gender issue. In these companies if gender plays

a role, then it is alongside psychographic or needs-based segmentations. Oftentimes, these adept brands address their customers on the individual level, with few pre-conceived notions of any sort. From the outside it seems to us that brands like Apple, Google and IKEA dialogue with their consumers in this way. With not a genderless but, perhaps, a trans-gender approach. (Note, such an approach is not to be mistaken for the brand itself having a gendered brand personality- a different brand aspect altogether.) As one MD said to us, "It's more about utility as opposed to gender— it goes beyond gender. I look at brands that are highly successful and they don't necessarily need to go one way or the other. Look at Muji or Uniqlo they are very egalitarian…I look at Gen Z- their ideas on gender are very different. They don't really define things by gender. It's very gender neutral."

Phase 1.2: Online qualitative consumer bulletin boards

As challenging as it was to reveal unconscious bias in businesses and brands, it was important that we examined consumers in the same way, especially since as previously noted, undisclosed prejudice does exist in society. This is what we did in Phase 1.2 through a series of online qualitative bulletin boards, where we collaborated with consumers of both genders from across Australia over the course of 3-5 days. Groups were split by gender and life stage, with a male and a female group for each category. Using narrative, 'show and tell' and task-oriented and leveraging a range of projective exercises and specialised elicitation techniques we explored consumers' perceptions of gender bias in the wider world as well as in relation to brand and business.

The flow of the discussion was guided by the understanding that we are all normed to the current status quo where gender bias exists. Hence the first day of interaction was dedicated to reflection upon how well their needs were being met by brands through communications, product design, and customer interaction, etc. in general, and for the categories of interest as a focus. Notably, we made no mention of gender at this stage. Respondents then participated in various storytelling exercises to help us understand where they felt their needs were and weren't being met, and why or why not. Throughout we listened out for spontaneous mention of gender bias or scenarios that could be interpreted as such. We then asked them to reflect on the scenario(s) they shared through a projective exercise. Using validated photo sets they were asked to select a photo from one of six photosets — respondents were only offered photosets of the opposite gender —retell their original scenario(s) through the eyes of that person including how it might have been be different for that person. In this way we hoped to elicit a gendered perspective.

On the second day of interaction, respondents were introduced to the notion of gender bias. They read online materials discussing gender bias and completed an implicit association test (The Harvard Implicit Association Test) to help bring awareness of personal unconscious gender bias. With this greater awareness, respondents were then asked to review the scenarios they described the day before and reconsider whether or not they displayed any gender bias and if so how. This allowed us to further explore manifestations of unconscious gender bias. They then were asked to describe what needed to be different to eliminate any gender bias and make it an ideal experience.

On day three, to gain yet more perspective, respondents were asked to visit a category retail outlet to experience in real life an aspect of what was discussed earlier. Respondents shared their experiences in detail and analysed them to better understand where and how gender bias manifests. Finally, respondents participated in a broader discussion of gender bias in daily life. Here, respondents helped us to reflect on the mechanisms, systems and structures that obscure bias and maintain the status quo.

The intuitions

From Phase 1.2 we were better able to understand the perspectives of consumers when it came to gender bias in relation to brands and business.

Like the marketers we spoke to, Australian consumers fell into the same broad typologies of awareness and understanding. However as consumers, and not active creators of products and brands, consumers are represented only by five of the nine typologies: Oblivious, Ignorant, Minimising, Cynical and Lost.

Consumer beliefs and attitudes that are barriers to change:
Consumers who do not acknowledge gender bias seem to do so for a few of reasons:

1. They are blind to the issue and content with the status quo: *Oblivious*
2. They do not recognise continued gender bias. They feel the issue is a 'thing of the past' or that it is being adequately addressed: *Ignorant*

Of consumers who do acknowledge it, they bear one of two attitudes towards it:

1. They see gender bias but don't think it is a 'big deal': *Minimising*
2. They are apt to believe that brands that actively addressed gender issues were not doing so authentically. (As a consumer typology this represents slightly different attitudes to the business typology): *Cynical*
3. They acknowledge gender bias but don't understand it- where and how it manifests: *Lost*

Many Australian consumers also had difficulty seeing past the historical foundation of gender bias, like the marketers we interviewed. For them, it was hard to realise that 'things can be different' when the status quo is so pervasive and persistent. When we first engaged with consumers on our qualitative bulletin boards, many initially found it difficult to highlight instances of gender bias in the brands and businesses around them. Many just simply did not see it; they exemplified the 'Oblivious' typology.

We found that younger respondents often, at first, rejected the idea of prejudice or bias on the basis of gender. They said, 'we aren't like that'. These consumers exemplified the 'Ignorant' typology. Perhaps initially their beliefs and values about the world somewhat clouded their ability to see the gender bias around them. However, once a few examples were shared, many of these respondents then were able to critically re-examine their worlds. As one male aged 18-25 years surmised "[It] made me reconsider the gender biases that are so embedded even in the products that we use that I probably never would've thought to question." Older male respondents were also well represented by this typology, however they were apt to challenge and reject showcased examples of gender bias.

As one 24-40 year-old man commented: "I'm not sure how much I agree with any of these videos, the image was a total and complete joke and while the article on women in the design field raises very valid points it also leaves out some important questions as to how they arrived at their opinions… Wow! Somebody is a real hothouse flower! How dare they put homeware in the same building as women's clothing, what sexist pigs they must be!"

Older male respondents also typified the 'Minimising' typology. With them highlighted instances of gender bias were often judged unimportant or inconsequential.

Not surprisingly, the 'Sceptical' typology found a place amongst older Australian respondents who were more versed in gender issues overall:

"#EqualFuture 'Pocket Money': great video, but makes a point that we are all glaringly aware of. I found it rich watching the video being made by ANZ having the first-hand experience with them at a corporate level. They now have 3 women on the board which they have used to push their credentials as an equal opportunity employer. Give me a break. Many studies now show women are increasingly becoming the predominant financial decision makers in the household and this is a clear ploy to appeal to that market. If ANZ were serious about promoting women they would be creating increased education and pathways for returning to employment for women and not just making feel good videos. The kids were cute though..."
Male 25-40yrs

Nevertheless, many respondents acknowledged gender bias once it was highlighted. They were then able to recall examples from their own experiences with brands and businesses. The category where gender bias was initially most apparent to them was often in children's products.

"I had a similar experience when I went to shop for my daughter in a children's store. I was looking for a blue or yellow dress for her on this occasion and do you think I could find one? The entire store was filled with pink, white or grey items for girls. I was so annoyed that I took a photo (which basically showed a sea of pink) and posted on social media to friends. The image is below for you to see. It just makes you rethink shopping there. Because apparently girls should only wear pink!!!" - Female 24-40yrs

Figure 1.

In a similar vein, as another female respondent 24-40 years recounted:

"I was shopping at Aldi with my son and daughter and had picked out two Kinder Surprises as a treat, for them being so good during the shop. I got to the register to pay for them and the cashier asked me if I wanted to swap one of the eggs because I'd picked up two girl ones. Up until that point, I didn't even realise that Kinder Surprises came in a girl version and a boy version. I was actually annoyed by the question because the cashier assumed that my son wouldn't play with girl's toys and vice versa. I really didn't care and told her thanks but no thanks. Why should certain toys only be for boys and certain toys only be for girls? My daughter should be allowed to play with the workshop toys, just as much as my son should be allowed to play with dolls, without others saying anything. These stereotypes are unfortunately being passed on to our kids from adults as such a young age."

Reflecting on how specific instances of gender bias with brands and business relates to the wider gender issue, one young man aged 18-25 years commented:

"Yesterday I went to my local McDonald's restaurant to grab some lunch. I did nothing particularly out of the ordinary - just ordered a McChicken meal, sat down and ate it. Having been part of this study however, I suppose I began to notice certain things about gender bias - most of the workers were girls. This didn't at all affect my experience, but I guess it might be reinforcing the stereotype of hospitality being a "woman's domain". I also noticed that the toys in the happy meals on offer were gendered - there were clear divides between "boys toys" which are currently the mechanical, strong, weaponised "transformers" compared to "girls toys" which are the "my little pony" range -pink, cute, helpless horses. It makes me wonder whether this is the beginning of where the male and female psyches begin to systematically change, and gender bias infiltrates our lives."

Gender bias in consumer touchpoints

Not only was it challenging for respondents to see the world through a different lens, with a different perspective, but they were often not sure if what they were experiencing was isolated and individual, or more universal and shared. For respondents to unequivocally identify incidences of gender bias, it oftentimes required them to share their experiences and realise they weren't the only ones facing such situations. Another phenomenon that made it tricky for respondents to identify bias was its oftentimes subtle nature.

Australian consumers seemed to more easily recognise gender bias in customer service touchpoints. For example, as one woman aged 25-40 years recounted, "There were one or two men before me and they also ordered coffee. When it came to my turn, I ordered a latte and was then asked if I wanted skim milk. The men before me were not asked the same question, so I felt it was possibly a bit gender biased." A woman 25-40yrs noted:

"I did notice the JB Hi-Fi shop was branded in a male dominated sort of way. Bright bold yellow and large fonts just seemed quite male orientated. I observed a lady being ignored at the computer section when she was browsing but as soon as a guy started nosing around, one of the sales guys approached him and said, 'how's it going mate' and struck up a conversation. I also noticed that on several occasions sales assistants didn't seem that happy that I was browsing in their shops with a pram and toddler."

Note: we hypothesise that customer service, being an interactive and dynamic touchpoint, may be more frequently reported due to availability bias, where human interactions are more prominent and more easily judged than, for example, having the ability to determine that a product or in-store environment could be designed differently to better meet the needs of a gender. This might require a better understanding for industrial design and engineering. Similarly, not having provided respondents with communications or advertising stimuli we suspect instances of gender bias for these touchpoints have been underreported.

Nevertheless, respondents did note that they were tired of condescending communications and pigeonholing. "The main issue with any kind of product's ability to serve women is usually that it's not marketed to women as actual human beings with individual needs and rather as stereotypical entities." - Female 24-40yrs.

Instances of gender bias in products were pointed out as well. A young man highlighted the 'male default' when it came to serving sizes, noting that his female dining companions did not finish their meals. Male respondents felt gender bias too. One young man felt uncomfortable asking the all-female staff at the skin care counter at his local department store for product advice for his acne.

Another young male commented:

"A recent gender biased shopping experience I had was going to shop at kikki.K. I've been there more than 10 times this year to go shopping with my girlfriend, and every single time I've only ever seen female staff. It sounds stupid to say this, but I also felt like the colour scheme of the shop was rather feminine - I know that colours shouldn't be gendered, but having pink and white and red hues in a shop staffed completely by females selling what is considered a traditionally female product - stationery - doesn't open the store up to male clientele."

When it came to instore sales environments touchpoints, respondents were more apt to note that they found aspects unappealing but often did not readily connect their complaints or discomfort to a gender bias. However, these can be seen as evidence of catering better to one gender's preferences and thus indicative of gender bias.

"I think the ideal store environment (McCafe) would have been calming and less frantic for the peak hour time of morning. I think dim mood lighting and some nice couches would definitely inspire me more. Somewhere to sit, so I don't feel rushed out, maybe a corner in the layout with that atmosphere and insulated for noise (a quiet area). The current McDonald's chairs are not comfortable and present themselves as 'buy something, eat quickly and get out'. The ideal sales interaction would not feel like a process; the sales assistants would be more friendly and strike up a nice conversation. I am quite partial to a caramel frappe but I don't think McDonald's make them...maybe they should." - Female 25-40 years

"I guess the ideal store environment (McDonalds) would be clean, easy to navigate and welcoming. It would make me feel like there was efficient service and that the staff were well treated and happy to be there. It would have an energetic feeling…The space would be easy to navigate- e.g. clearly marked lines which mean you're not getting in anyone's way. There would be clear displays with all the menu items displayed. The prices of all items were not available during my visit. The sales interaction should be friendly and energetic without being over the top- a smile and a genuine 'have a nice day' is great." - Female 25-40 years

Female respondents in particular felt that stores did not adequately cater to their needs when they were accompanied by children. For example, self-checkout spaces were noted for being too small for those shopping with kids. Cleanliness also seemed to be a concern for females.

"I would have to say I hate shops in general, parking is difficult and isn't wide enough, isles are not wide enough, we want to serve ourselves because as a female I am independent but there's not enough room at self-service especially when you have kids with you. Some of the stores layouts make you wonder what they were thinking. I mean if it's a spread put it with the other spreads just cause it's golden syrup doesn't mean it should be with the cooking stuff. Have more people working- women are busy people and kids hate shops. Maybe have something to entertain kids while you shop. The trolleys really need to be cleaned, they need to be stocked well and some products are just up too high. I mean why make the shelves that high. We are female and are generally a little shorter. I like to discover new items but don't have the time. Why not have a special area designed for new items. If they changed these things I think shopping might not be so bad." - Female 25-40yrs

We can see more typically male preferences shared by this male respondent's reflections about electronics stores:

"*The ideal store would be quick service, I'm in and out of there at the fastest possible time. Ranking of prices would be more clear so I'm not spending 10 minutes trying to figure out the best deal A rating system of each product has performed from other customers themselves As a male i like my space simple and organized. What i hate about some stores is just finding an area. Why not make everything pre-visible instead of hiding an item in a corner where you have taken an hour to find it. Maybe a touch screen could be provided to show you where an item is stored, or if there is one there or not. There have been many occasions where I've been to a store to not find what in after. Wastes a lot of my time. You can ask an employee but sometimes they won't know, and have to check the computer themselves to find out. I should have that access. Perhaps an app on my phone? In terms of what the store looks like I'm not fussed. Just have it clean. Looked after. And enough space to move around. Items easily found. I don't my products to be glamorised. I just want to know its contents and any information useful to me and what it's for."* - Male 25-40yrs

Category gender bias

Gender bias was also noted in specific categories such as automotive and home repairs, and fashion as well as electronics, discount retailers, fitness, supermarkets, and stationary (as previously noted). As one female reported, "I think gender bias is extremely strong in the fitness industry. Women are meant to lose weight and men are meant to bulk up and be strong. Trying to go to the gym to be healthy and because it feels good and I want to feel strong and manage my mental health is so annoying when all the instructors

and media is all based on losing weight and becoming less. I would love the fitness industry to cater for individuals rather than stereotypes and accept the fact that people want different things and it's not because of a gender."

Rejecting female stereotypes and highlighting the deeper complexities of the gender issue, another woman relayed:

"I find shopping centres like Westfield inundate you as a female with images of the ideal woman. White, tall, slim and young. Makeup stores, clothes stores and jewellery all either have posters of this ideal woman around the stores or hire staff that fit this description. I'm an average sized woman (size 10/12) and feel intimidated when shopping because advertising like this makes you question whether the products they are selling you would suit you before you even walk through the doors. I find that the staff are not adequately trained so aren't attuned to what the customer wants or needs. I really think advertising needs to start with correct representation. This country is diverse culturally and women come in all shapes and sizes. People respond to relatable material. If advertising can get this right, women and men would be more inclined to purchase what is being sold."

Figure 2.

This is the associated image she shared.

Ultimately, Australian respondents recognised that the historical foundations of gender bias wouldn't change unless consciously addressed. Notably in our sample, younger respondents often highlighted gender bias towards the opposite gender.

We can see this in the comments of an 18-25 year old male who noted the sexualisation of females in communications:

"I went to a Korean fried chicken lunch bar in China town where the main theme marketing strategy seemed to be using cardboards cut-outs of lightly clad K-Pop singers that had speech bubbles about their favourite items on the menu…It seemed like both excessive and needless sexualisation of females to sell a product. I don't think companies need to rely on using females as a marketing tool to entice greater sales

from the male population. What happens is that male customers like myself feel irritated and uncomfortable about being at a place like this."

Phase 2.1: Quantitative survey

As of this writing, we have yet to complete the second and final phase of research which will be quantitative. In it, we will survey a nationally representative sample of n=1000 men and women in each of the two categories. Each category will also have a control cell of n=500 for a total of n=3000 respondents across both categories. The 2 x 500 respondents in the control cells across both categories will not be alerted to the gender intent of the research; the exposed groups will. This is a key design feature of our approach as it will allow us to understand and quantify unconscious bias or the impact that being made aware of the gender issue has on perceptions.

In the first section of the survey, we will present exposed cell respondents with priming vignettes. Participants will answer a series of questions in response. For example, they will be asked how these vignettes make them feel and whether or not they would respond if they were to witness something similar in real life. These form part of TNS's proprietary Political & Social ConversionModel 'States of Mind' framework and results will be used to establish individual respondent 'gender inclusion' baseline attitudes. Comparison between 'State of Mind' groupings will be used to deepen analysis of the primary outputs. More fundamentally, exposure to these vignettes will serve to prime respondents with a gender awareness which potentially will impact upon responses to the core set of survey questions.

All respondents n=3000 for both categories and in both the control and unexposed cells will then complete the core part of the quantitative survey, where we will quantify the equity of brands in each category, using TNS's traditional ConversionModel. We opted to use this framework because of its ability to reflect the real life decision making trade-offs people make in the competitive environment. Here, respondents will rate the brands they use or would consider using on their performance using a 1-10 scale where 1 means "Terrible" and 10 means "Perfect". Then, they will indicate on a 7-point scale to what extent they feel those same brands connect with whom they are and with the things in life that they really care about. Respondents in the exposed cells will respond to both of these same questions but with the additional, "as a man" or "as a women" verbiage. Alongside the priming vignettes, this additional wording will help to make the gender issue more salient. This will supply a holistic perspective of the category and its main brands by gender and provide context for touchpoint ratings.

Finally, respondents will rate each brand on nine touchpoints: TV advertising, outdoor advertising, social media, word of mouth, sponsorship activities, instore menu or product point of sale displays, instore retail environment, customers service and product design and usage. Respondents will rate brand touchpoints, for brands they use or consider using, for extent to which they feel it connects with who they are and with the things in life that they really care about. Again, respondents in the exposed cell will respond to the questions with the additional "as a man" or "as a women" wording.

With a complete picture of the brand landscape from both the business and consumer perspective we will be able to bring light to the subject of gender bias, revealing with greater specificity (qualitatively and quantitatively) where issues lie in categories, brands and touchpoints.

Ideally, this research would support a Phase 3, whereby we could link the quantitative consumer brand performance by gender of Phase 2.1 (external) to the organisational mindset qualitative typologies of Phase 1.1 (internal). This would enable organisations to understand how to transform the business internally first in order to resultantly impact on brand perceptions and performance in the market. This would require us to research and interview each brand organisation covered in our quantitative sample and unfortunately is out of scope for this inquiry. However, we will hypothesize which typology each brand fits into based on the quantitative results of Phase 2.1 in order to deliver a holistic perspective of business evolution.

Research importance

With the constant reminders of gender equality, we can forget that the most powerful influence on a consumer's buying behaviour is his or her gender. Businesses today still have much to learn about how to address the different needs of men and women. This study presents a challenge to the continued gender stereotyping and misunderstanding perpetuated by brands and businesses, who have stalled in their efforts to address the needs of the genders. In particular, despite significant progress, we hope to bring attention to how undervalued and underserved women are in the marketplace. Our work, here, builds on the theory and practice of other researchers and marketers. It does so qualitatively by continuing the dialogue with consumers today in Australia, and quantitatively by measuring their experience and quantifying the impact in dollars. Importantly, our work acknowledges and measures the impact of unconscious bias as a contributing force to the perpetuation of the status quo. Notably, our work also acknowledges and considers the dual sources of biases within businesses and consumers, alike.

As one woman we spoke with said about the future:

"It has gotten a lot better as I have [been able to move up the ranks and] have more control over what happens. So, I think the more women you have in leadership roles, the more there is going to be a recognition that women are not 'small, pink men', and that our needs and wants and desires and drivers for what we purchase are [different]. I hope that the products become more aligned to what we need. I hope that the communications become more what we want and need. I do have hope and faith, but we have a long way to go."

Today, businesses increasingly realise that they have the opportunity and, perhaps, the responsibility to impart a positive impact on society, not just from an ethical point of view but also a practical one by making things that both genders can use and enjoy. By re-examining how they serve the two biggest segments of consumers – men and women – brands can uncover areas ripe for improvement, so that they make a change for the better.

We believe that brands that ignore this movement not only relinquish growth to competitors who do, but they risk being left behind. And if you think it is a zero sum game – think again. A female respondent summed it up nicely when she remarked: "So marketing to both sexes will bring the market and $$$ up on products that feature both men and woman." When done well, increasing profits with women does not have to mean sacrificing share with men. It can be a win all around.

Perhaps, future generations may not have to wrestle with this issue, but for the time being a base of underlying gender bias obscures real needs and satisfactions among men and women. However, as market researchers, we have to acknowledge our contribution to the problem. We, too, have been complicit – we haven't availed our clients of approaches or tools to examine these issues. Here, we offer one. It's a start.

Acknowledgement

The authors would like to acknowledge and thank the following people and organisations: Anne Rayner for her leadership and direction in initiating this line of inquiry. Lesley van der Walt, Mike Griffiths, Kathy O'Donoghue, Jon OLoughlin, Alison Dexter, Eleanore Wells, Brian Walker-Catchpole, Apple dela Rosa and Angela Hoffman for their selfless contributions of time, effort and expertise. Ava Research, CiviCom and Lightspeed Research for their support of this project.

The Authors

Amy Fridlund is Insights Curator APAC Brand & Shopper, Kantar TNS.

Nikki Feld is Senior Consultant, Kantar TNS, Australia

PART 7
Making Big Discoveries: Data and Analytics Capsule

Integrating Survey Data and User Data

Bridging the gap between “Who I think I am” and “Who I actually am”

Yang Yang • Bonnâe Ogunlade

Introduction

In the age of performance and programmatic marketing, advertisers have come to expect speed, agility and accuracy in everything from target audience understanding, to communications planning, and measuring campaign effectiveness.

To meet the challenges of media planning in a convergent world, Carat has been investing in a proprietary consumer and media study, the Consumer Connections System (CCS), for a decade. Now with 133,664 respondents in APAC alone, these nationally representative surveys of each market comprises of questions about consumer demographics, lifestyle interests, attitudes and values, media consumption and behaviour across 60+ touchpoints, advertising receptivity, as well as product purchase intent and usage.

As our clients use CCS to define and understand their most valuable consumers, we wanted to improve our offering by finding ways to integrate declaration-based insights from survey data with the near real-time behavioural insights from other sources.

The three key benefits are as follows.

1. Insights

Digital behaviour, such as keyword searches, clicks on ads, and visits to particular websites are used to classify people into affinity segments or in-market audiences. Whilst data collected from CCS helps us to understand the profile of consumers who have claimed purchase intent for a certain product, we have the potential to identify people who are further along in the consumer journey as demonstrated through their digital behaviour and understand more about the types of content they consume, and the channels they turn to.

2. Media Planning

Equally important is the need to determine and verify the media channels most effective for engaging with our target audience. Media surveys have long relied on consumers’ claimed exposure to media and behaviour on each channel. At the same time, there is a proliferation of data sources that measure and report, in near real-time the exact audience interactions with media touchpoints, such as television audience measurements and digital audience monitoring. The accuracy of media survey results can be augmented with such data.

3. Activation

With data-driven understanding of the consumers passing through their conversion funnel, advertisers will be able to deliver more relevant, even personalized, promotional messaging about their products in the moments that matter. Consumers’ data, whether declaration based or captured via real behaviour, is to be used only in good conscience with the end result of the data connection benefiting the consumer, offering them utility over and above mindless targeting.

In this paper, we share details about three data integration projects that Carat, alongside our network group Dentsu Aegis Network, have undertaken in the APAC region, as well as a concept soon to be explored. Each case study is explained through the context from which the need for the project emerged, the data sources involved, and examples of the benefits delivered to clients.

Case study 1: Augmenting media survey data with reach curves

Background

Dentsu Aegis Network is devoted to developing a people-based marketing proposition and the Consumer Connection System (CCS) is our system for understanding people and delivering enhanced insights. The system is centred on an in-depth survey that covers demographics, motivations and media.

Additionally, our single source survey data powers our convergence planning tool CCS Planner, which enables users to map insights to media behaviour. The fact that it is single source means that users are able to plan across multiple touch points while accounting for duplication.

Objective

Our desire is to successfully engage consumers on behalf of our clients by having a strong command of the current media environment. However, this can be challenging due to the abundance of media sources and channels and consumers' increasingly convergent media behaviour. To achieve the most effective planning results, we believe that it is critical that all users validate their plans using the most up-to-date data sources available.

CCS Planner gives us our foundation by providing a multimedia view of consumers' media consumption habits. The current system is refreshed on an annual basis using the latest CCS survey results. Accordingly, CCS Planner is fully enabled with the functionality to calibrate key inputs such as costs and reach data.

In conjunction with the calibration function, we are devising an approach to overcome such obstacles, specifically in markets where behavioural data sources are limited. We are looking to deploy a shorter more agile mobile survey to refresh the fast moving survey data such as media consumption habits and device ownership.

Methodology and findings

The common approach for calibrating reach data is via third party data sources such as the digital measurement data provided by ComScore and TV viewing data available via Nielsen's Television Audience Measurement (TAM).

Connectedly, the growth in digital video consumption means that an accurate solution to capture viewership and truly understand the value of online video is also required. Hence, why as a network we have acquired exclusive access to reach curve tools for YouTube and Facebook. The partnership gives us access to YouTube reach curves in 32 markets and Facebook reach curves in 17 markets with the aim to roll out access across our 60+ CCS markets.

To foster maximum usage, these tools are assimilated into the CCS Planner platform making them immediately accessible to all. This means that every time a user creates a plan that features either YouTube or Facebook platforms, they can refresh their curve data with the most up to date and accurate reach information, which ultimately provides a unique, consistent and altogether more robust solution across markets.

We continually test, learn and measure the media that we plan so as to optimise our recommendations and share best practices internally and with clients.

Our measurement is often in the form of econometric studies that allow our teams to review the business value of their media plans. Specifically, they evaluate Return on Investment (ROI) and inform optimal allocation of the marketing budget across channels.

Case study 2: Integrating survey panels and TV audience panels

Background

We systematize our planning process through our in-house convergence tool as it allows users to plan against the clients' target audience, campaign objectives and Key Performance Indicators (KPIs). Yet, the continuing stagnation of the economic climate means that our clients require more accountability, accuracy and greater efficiencies from our media buys.

Objective

Using data fusion techniques we are able to integrate two powerful data sets to not only identify our clients' most valuable consumers, but to also activate media buys against the same target audience via the most suitable platforms and devices.

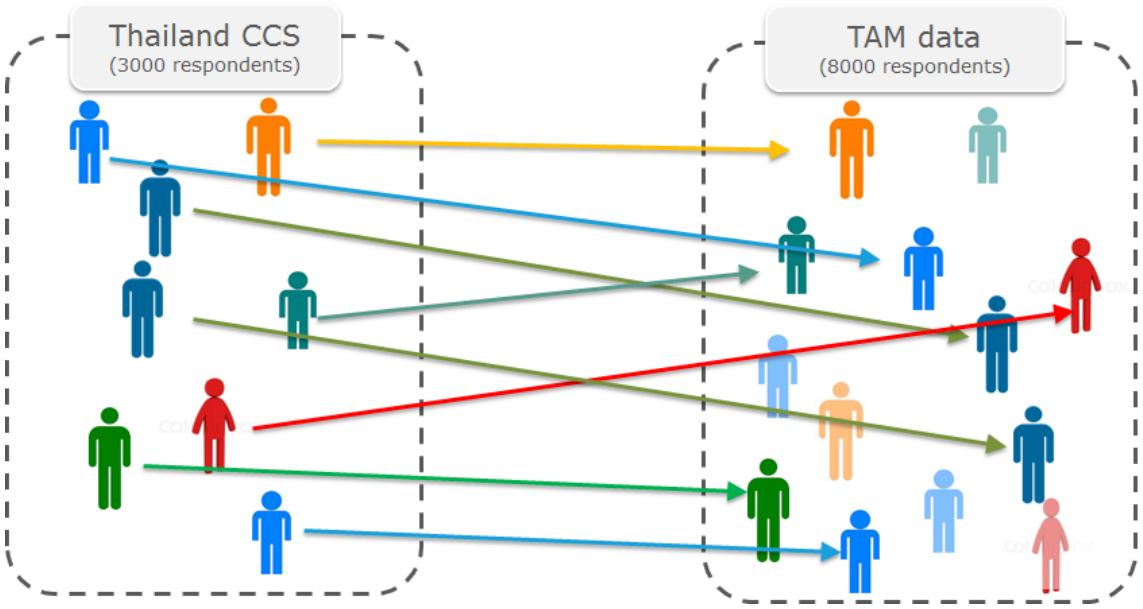
Methodology and findings

In Thailand, we have fused our consumer survey data to the market's television audience measurement data to deliver greater buying efficiencies for TV spots.

Data fusion is based on a statistical technique that: 1) identifies linking variables between two respondent level data sets, and then 2) leverages analytics and modelling to create a “single” data set containing the attributes from two or more data sources.

To create the strongest data set, our consumer survey acts as the donor dataset (the dataset that will give/donate additional variables) to the TV data set. This is because the TV data is more dynamic, as it gets updated more frequently, making it a better candidate to be the recipient dataset (the database that acts as the main database with the donated additional variables).

Figure 1.

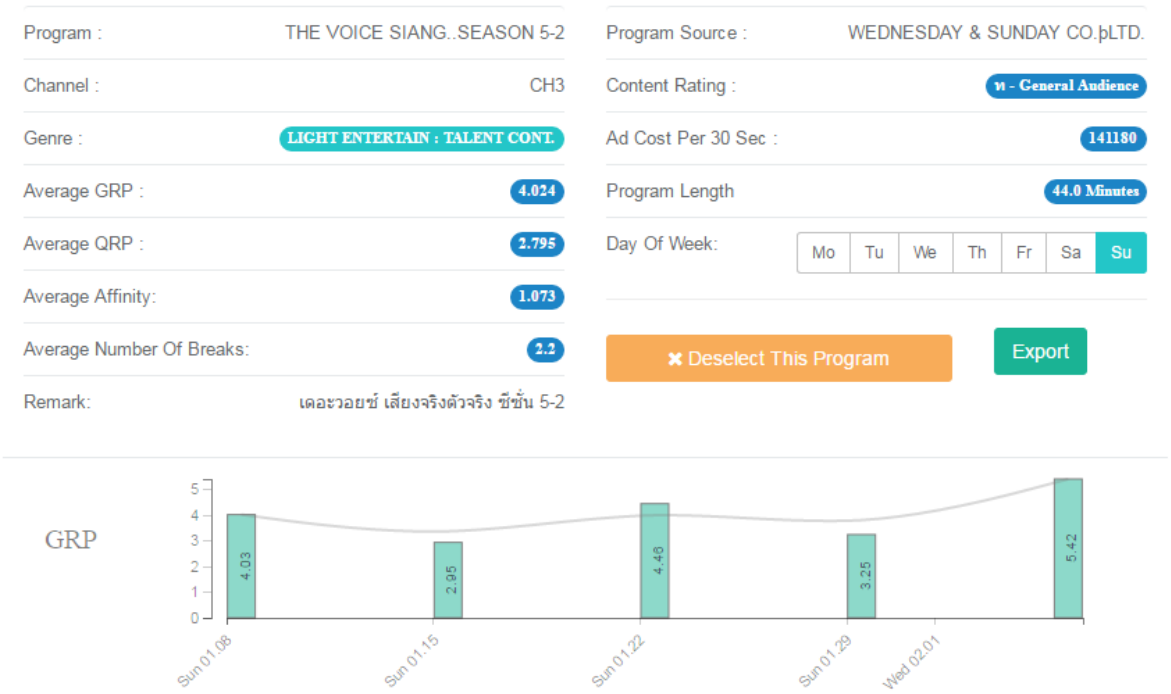


The overview:

- We have established that in this case our proprietary data is the donor data, and TAM is the recipient data. For every person in the recipient data (TAM), the fusion technique will select the person with the most similar profile from the donor data (CCS), using a number of common variables
- Neither the donor data nor the recipient data hold any personally identifiable information.
- The final outcome is that the fused data contains the TAM data from the TAM respondents, with survey data appended to each TAM respondent

- We can define in our TV planning software the same audience that we have defined in CCS and we can analyse the audience against trading audiences
- To further enhance our fused data, it has been loaded into a complementary proprietary tool. The tool's purpose is to calculate the attention and engagement levels of viewers for a particular TV programme, therefore assessing the quality of the TV viewing audience of a particular TV spot

Figure 2. Program detail

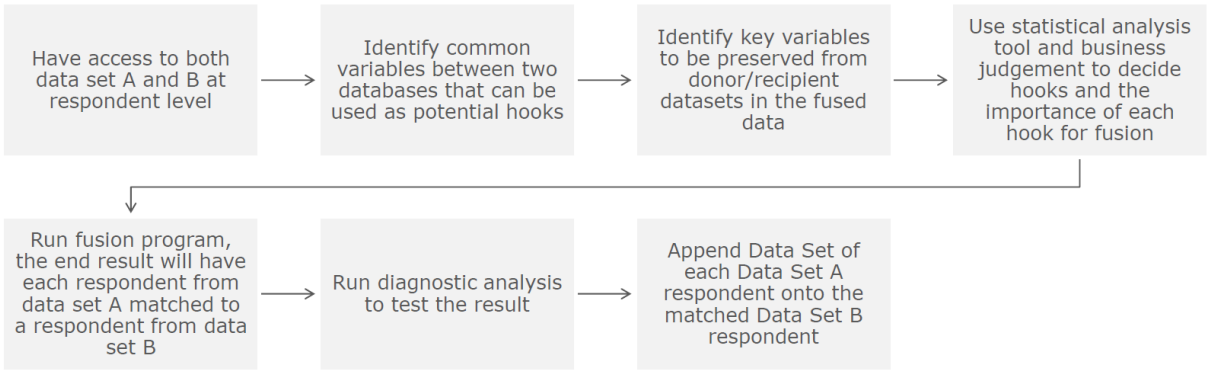


(We have developed a tool to empower our TV planners and buyers)

How to make fusion happen?

The success of a fused study relies on the quality of the databases and the common variables that are used as the linking bridges to connect them together.

Figure 2.



Case study 3: Connecting datasets via a proxy panel

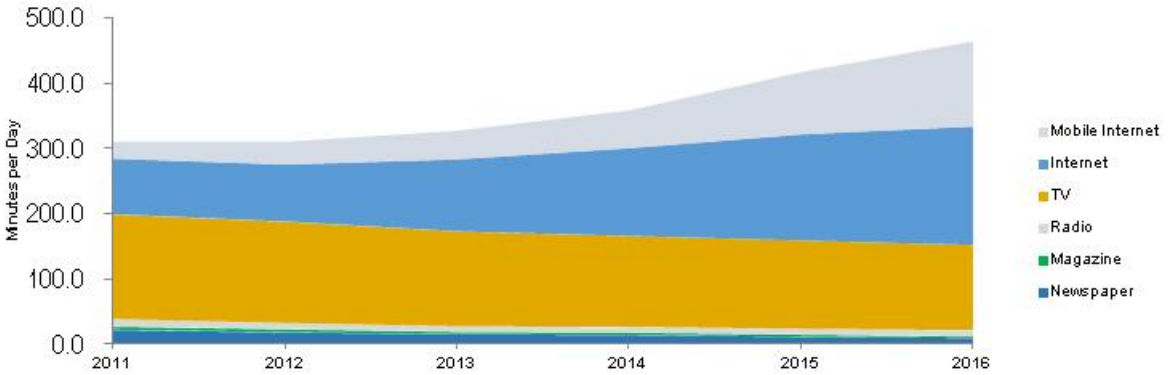
Background

CCS China is one of our largest studies in the world, with 70,000 respondents across 88 Tier 1 to Tier 5 cities.

A major shift in consumer behaviour that we have observed over the years was in the amount of time spent on mobile phones, excluding the time spent making calls. Between 2011 and 2016, time spent has

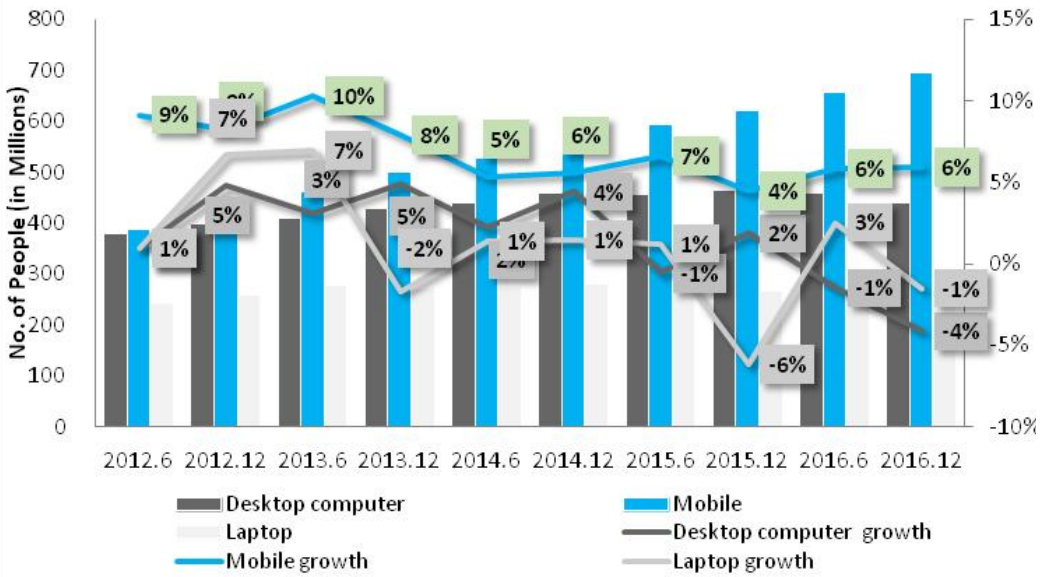
increased five-fold from 26mins to 130mins. In fact, mobile has overtaken computers as the leading platform for accessing the Internet. We needed to understand more about how this time is being spent so as to pave the way for mobile-first media planning.

Figure 3. Time spent on media



(Average daily time spent on each media channel by Chinese individuals. Source: CCS China 2014-2016)

Figure 4. Desktop vs. mobile internet users



(Volume of people accessing the Internet via desktop and mobile devices. Source: CCS China 2014-2016)

Enter Google, Carat's global media partner who is keen to break into the China market. Google's Android OS is popular within the market and its analytics and advertising services continue to be operational. This means that there exists a wealth of information about mobile users, such as the advertisements they have clicked on, apps accessed, length of stay, etc. However, the key challenge faced in mining this data for consumer insights is that while each string of mobile activity is mapped to a cookie, we do not know much else about the person behind the ID – what is their gender? How old are they? What products do they intend to purchase, and what are their attitudes towards advertisers?

By bringing the two sources of information together, we could learn more about the types of mobile content and experience desired by the key target audiences of our clients which are being defined in CCS.

Objective

The first challenge we faced was finding a way to merge the CCS dataset with Google's. This dataset does not capture any identifying information about its respondents, such as name, contact number or email addresses. As a result, there were no common variables between the two datasets that we could use to 'hook' them together.

Since matching CCS respondents to their Google identity was out of the question, we need to establish a logical framework for data fusion.

Methodology

The only way forward for fusing the two datasets was to first create some common variables between the two. We did this by establishing a proxy panel of respondents amongst Android mobile users. Using Google admob, we fielded e-invitations to a survey in the form of text ads, standard image ads, and interstitial ads.

People who clicked through these ads on their mobile phones were taken to Carat's mobile survey platform with questions covering:

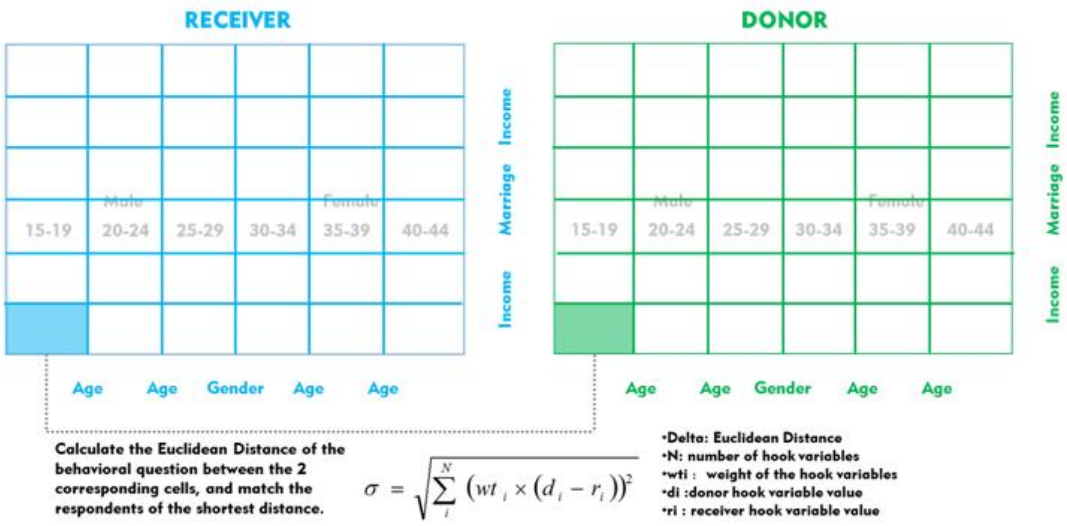
1. Gender
2. Age
3. Marital status
4. Age of children
5. Personal Income
6. Type of content browsed on mobile

We also included a couple of questions pertaining to ownership of products relevant to the advertisers we work with so as to facilitate identification of target audiences later on.

As Google had advised that the average click through rate of the types of ads we are using for recruitment was 0.1%, and as we knew that survey completion rates on our mobile survey platform was 5%, we planned for 200 million ad impressions so as to achieve a target completion of 10,000 surveys. Throughout this process, we tracked the ad response through DoubleClick Campaign Manager and eventually a sample size of 6,011.

Google's user data on the proxy panel was then processed through a data fusion software with CCS respondent level data. The premise of this fusion process is that we were looking for people from both datasets with the shortest Euclidean distance on a key behaviour – type of content browsed on mobile – who are from the same demographic cells.

Figure 5. Desktop vs. mobile internet users

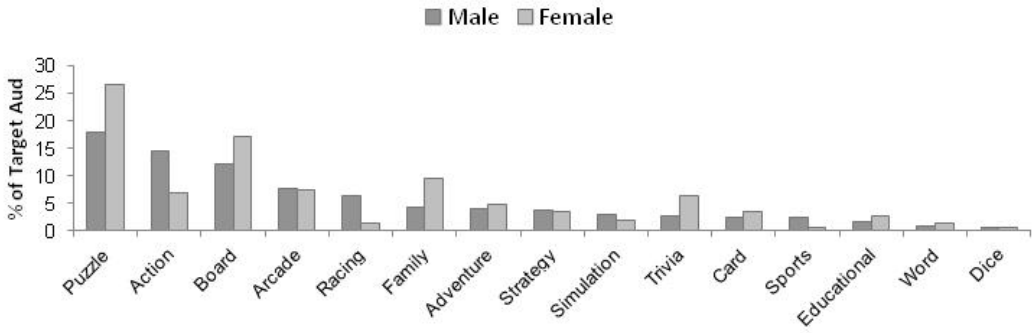


Findings

In CCS, we ask about the genre of apps that our respondents would regularly access, such as weather, utility, gaming, etc. However, due to limitations with the length of interview, we could not delve deeper into the specific sub-categories (e.g. type of games played), or put these insights through useful filters such as day of week, time of day, etc.

Through the Google Mobile Study, we were able to link target audiences with their real behaviour. For instance, we uncovered that between the male and female target groups of one of our clients, females played more strategy games such as Puzzles or Board games, whilst males preferred animated and fantastical games like Racing, Simulations, and Sports. As a result, we are able to advise advertisers about the specific gaming apps to advertise on for products skewing towards either gender, and also propose partnership opportunities.

Figure 6.



Acknowledgements

The authors gratefully acknowledge the contribution of Simon Zhong, Senior Analytics Director; Angela Chen, Group Planning Director; and Jane Chen, Research Director (Dentsu Aegis Network, China); Jaratpan Onghununtakul, Strategy & Research Director (Dentsu Aegis Network, Thailand); and Derek Huang, Head of Insights & Research, Silvia Leiva, Director of Insights, Derek Luo, Senior Analytics Manager (Dentsu Aegis Network, APAC).

The Authors

Bonnie Ogunlade is Regional Insights Director, Dentsu Aegis Network, APAC, Singapore.

Yang Yang is Regional Insights Director, Dentsu Aegis Network, APAC, Singapore.

.

How Can a Chinese Manufacturer Discover their International Customers through Integrated Methodologies?

Creating a lens into international markets by layering technology-enabled research to discover deeper, richer insights

Jeff Tsui • Alex Tu

The background

For many Chinese manufacturers, their target consumers are based internationally so accessing an accurate view of their audience is arguably more imperative than those who are on the ground and living in the same culture as their market.

As Chervon expands further into the US and Canada with the plan for the first battery powered ride on lawn mower, the need for consumer data across as many metrics in the form of tangible, actionable insights was identified. The geographical disconnect is the case for many Chinese manufacturers. As trade agreements face tension under new US presidency claims, it has never been more imperative to truly understand their audience and maximise their market share for the future.

At the same time, technology, by its very nature, is making it easier to enhance insights by tapping into what consumers do, not just what they feel. Having the technology is one thing but leveraging this for true impact is about fusing these technological advances into our experience as researchers and creating a greater output to drive the insights we offer brands for their future.

The phrase ‘technology is making the world a smaller place’ rung true to us here. Although it is normally around advances in transportation and communication, Lightspeed and Chervon set out to see if the same applies to market research.

The objectives

- Explore if research provides a window to the world through different lenses that combine into one research solution and data set?
- Discover how new tools can add layers of detail and open the door to deeper and often new insights for researchers
- Share findings from first-hand experience from the project and how this impacted Chervons ability to progress their international markets
- Show the risks and rewards of exploring new technology add-ons for your research

The methodology

Engaging Lightspeed, the two created a lens into these countries by layering technology-enabled research to discover deeper, richer insights.

The starting point was a standard 15 minute mobile quant survey. The study was conducted in late Jan / early Feb 2017, with N=1,041 respondents across US & Canada. This was targeted at respondents +30 years old and screened for either P2Y ride on lawn mower purchasers or those intended to buy a lawn mower in N1Y.

To develop this new integrated hybrid qual-quant approach the survey was overlaid with:

- *Video responses* - participants record and upload a short video to answer three OE questions within the survey rather than typing their answers, delivering video reel and transcriptions
- *Social listening data* - using a global social listening platform to search keywords and gather what customers, prospects and fans are mentioning
- *Facial coding analysis* - respondents record their face as they watch a TV advertisement to measure the emotional effectiveness of an advertisement.
- *Data appends* - combining consumer third party data with panellist profiles, this marries behavioural and attitudinal data to paint a more complete portrait of the target consumer.

By building four levels of further data this paper highlights how today's technology can give much richer, deeper discoveries and how Chervon were able to use this for a more effective approach to growing international markets.

The findings

Ultimately, the integrated hybrid qual-quant approach helped overcome the geographical gap and also developed a deeper understanding of the audience that delivered value and impact on how they should proceed into international markets with a new product concept.

Quant: Preliminary survey findings

The who

As found from the survey data, amongst those who have purchased a ride on lawn mower in the past two years, we recorded an almost even spread between gender, and no significant differences between those age groups was observed. Amongst buyers, there is a slight skew towards those with annual HHI of US$100k or above. In terms of riding mower type purchased, 67% bought the front engine version. We therefore believe the core demographic segments Chervon should focus on are those aged 30 years or above (i.e. from new home starter) with an annual HHI of US$100k or above; and it is important we do not stereotype and focus marketing effort towards males only. Aside from targeting the group that is already considering a ride on lawn mower / to gain market share within the same segment, we also believe strongly that we can convince some of the push on lawn mower buyers to up trade to a ride on lawn mower, as 87% of them claim that they would consider a riding mower powered by a battery if available.

Current usage, purchase drivers, preference and brand usage

From the survey, respondents told us that they are looking for convenience when it comes to a ride on mower. The most important attributes our respondents look for when mowing their lawn are "quality clean out" (77%) and "get it done at once" (62%). 47% also rated "ease of maintenance" and "good after sales service" amongst their top three important attributes when choosing a ride on mower. Most importantly, they are looking for a quality product that will do the job they set out to do, which is to get a good lawn after mowing it. Over 50% of the respondents rated this as the most important attribute.

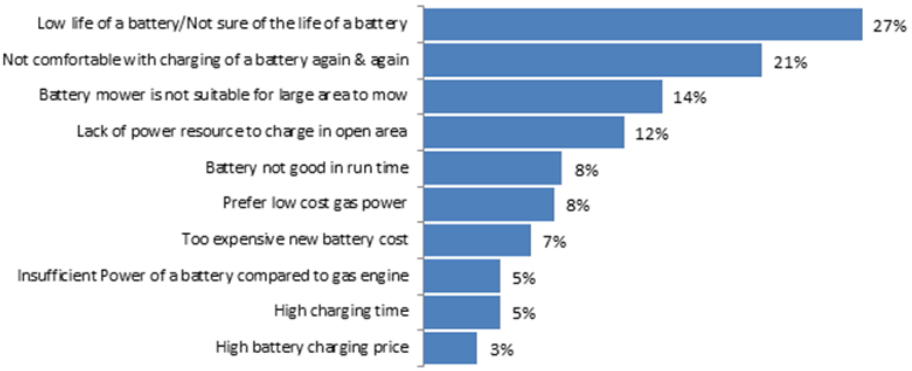
Testing of new product concept

When asked would one consider a ride on mower powered by battery similar to an electric car, 77% from US said yes, and 88% from Canada said yes. There is also a clear trend to suggest that the younger the age group, the higher the acceptance level of such a concept.

Features and functionalities

By asking respondents who claimed they would not consider a battery powered riding mower, the top concerns were found to be "battery life and run time" and "unsuitable for large area to mow". These deliver a list of functional requirements key for the development team to further explore in the final product.

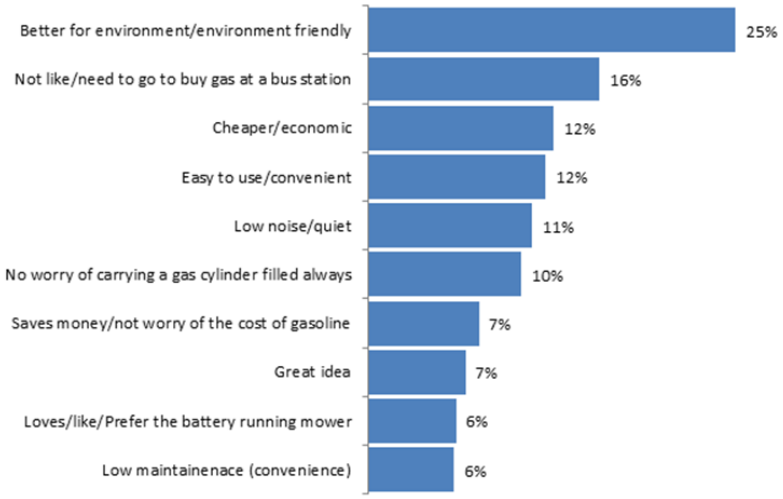
Figure 1. Barriers to consideration of a battery powered ride-on lawn mower (Top 10)



Hints at positioning ideas from open ends

On the opposite end, the potential benefits respondents envisaged for a battery powered ride on lawn mower are lower noise level; less smelly; most cost effective; less effort in maintenance; higher durability; more stability; smaller in size; safer etc.

Figure 2. Benefits of a battery powered ride-on lawn mower (Top 10)

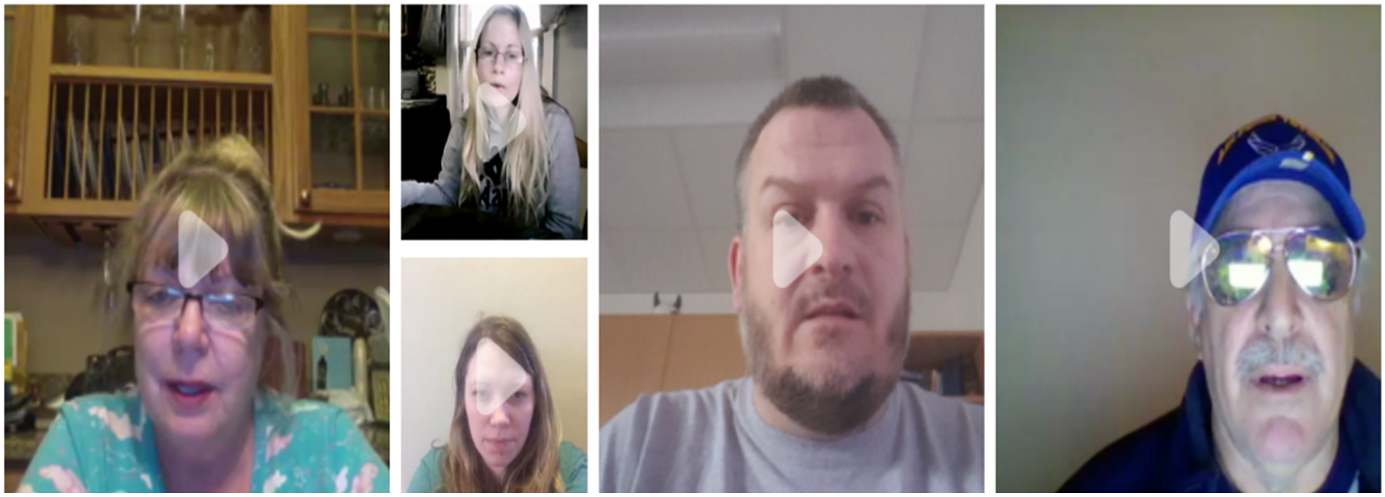


Aside from the obvious functional benefits, there were also female respondents who look for fashionable design and a degree of customization (e.g. colour of the machine). Pricing is also an important consideration; the survey found that respondents spent an average of US$2,000 when they purchased their last ride on lawn mower.

The survey gave guiding data to build from; however this was just the start with four more layers of data to be incorporated to this first phase of market research.

Video responses: Diving deeper

Figure 3. Video stills

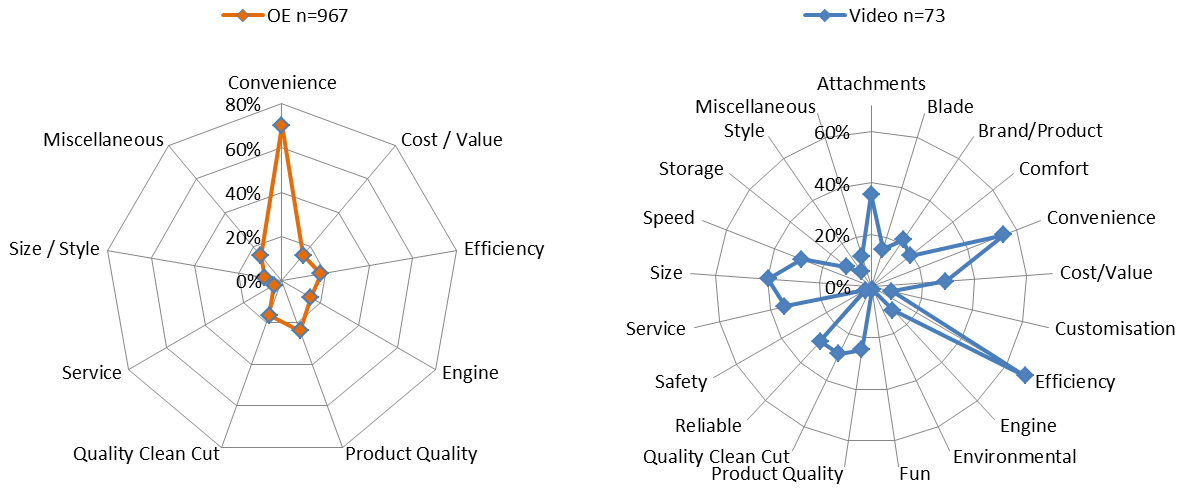


Reasons + feelings = Enhanced way

Offering respondents the option to answer open ended questions opened up a far more naturally flowing and detailed explanation behind some decisions reported in standard survey choices, including pride and prestige which seemed important to many;

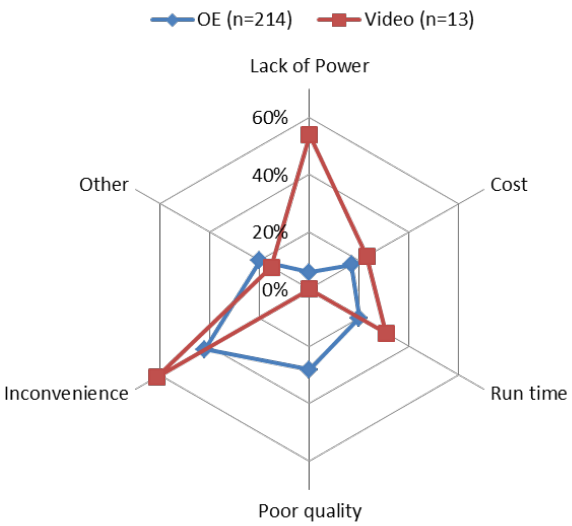
- *"Something that would look presentable as far as when I'm riding outside. Neighbors see me riding it, something to be proud of."*
- *"John Deere… probably the Harley Davidson of riding lawnmowers"*

Figure 4. Appeal of the ride on mower: Open ends vs. video responses



Overall, video responses contained 76 words vs. Open ends which only contained 18 words, giving Chervon more data to understand their audience, a qualitative element that would traditionally involve in person focus groups and therefore a trip to the market.

Figure 5. Barriers to ride on mower – Open ends vs. video



Product enhancement cues

Video responses both verified and added to the findings of what potential purchasers would be looking for in this product, key data for Chervon to have before development even begins.

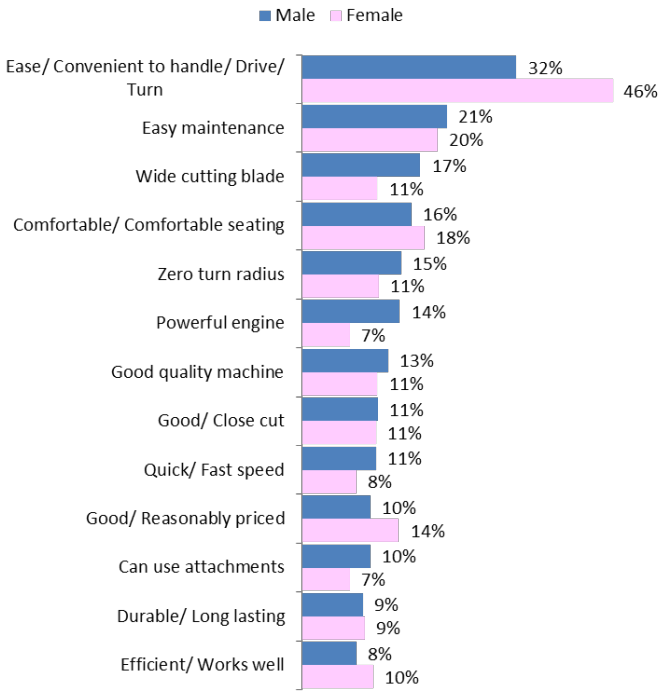
The survey data had already shown the want for mowers to be convenient but now we found this also included how to store them;

- *"…light and easy for me to use as a single woman. Something that I could store easily."*

Female requirements seemed to grow here, these prospective users want the mowers to be smaller in size so that it's easier to manoeuver, as well as the controls so that they are easier to handle;

- *"My dream riding mower would be pink, it would be powerful, it would be compact so it could easily fit into the garage and I'd have some safety triggers on it of course for the kids not to be able to just turn it on and use it whenever they felt like it, yeah."*

Figure 6. Ride on mower features – Male vs. female



What also started to appear is a similarity between cars more so than standard DIY machinery. This is understandable given the focus on a ride on mower but this moves Chervon from their comfort zone and made the delivery of layers of data beyond the standard approach even more critical. The idea of the mower being an extension of one's self and a status symbol was surfaced by video responses. So too was a newer but hot topic for the automotive industry- environmentally friendly:

- "Battery operated is good idea because it's environmentally friendly"
- "Using electricity just seems like a clean choice…"
- "..so economically and environmentally, I think it would be responsible.."

And the enjoyment and comfort of the ride:

- "…a smooth luxurious, high quality ride with a radio."

This hybrid study had been able to uncover a hybrid category for the new product concept. With the rise of Tesla in the US in particular, Chervon were able to see huge potential in the market.

Social listening: The bigger picture

The next layer moved away from the mode of asking questions that we as researchers are so used to following and instead listened to what was already being said across social media.

Conversation starter or not?

Initial searches for "ride on mowers" and "electric ride on mowers" returned low mentions. However, this comes as no surprise given that the electric ride on mower is no more than a concept at this point in time. More generally, ride on mowers appeal to a niche audience who therefore do not talk so broadly about their usage or desires across social media currently. This is to be expected where incidence rates are lower.

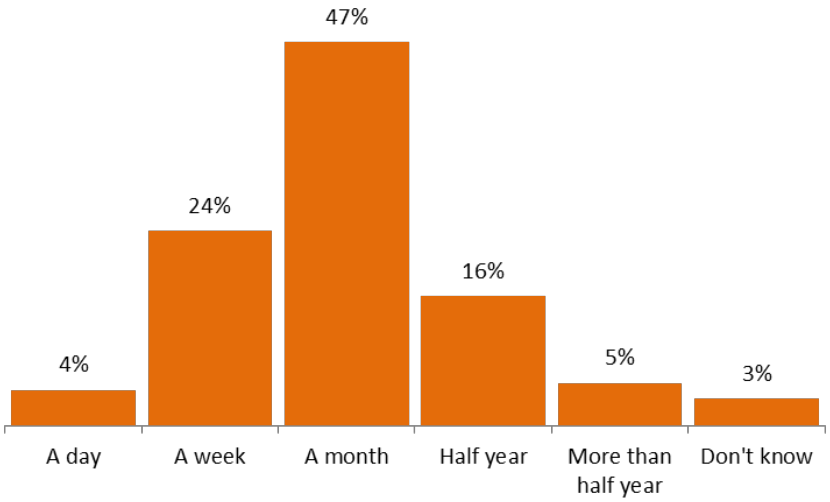
It could be read as social being a failed data source at this point but in fact this lack of mentions is very insightful in itself.

Marketing and advertising channels

In fact, the lack of presence of this product type and therefore lack of competition within the market and within social media channels gives Chervon the chance to broaden its own exposure.

This may help narrow the segment, compared to expensive options with potential wide reach such as TV ads. The survey data suggested that 75% of lawn mower customer's research intended product purchases within one month prior to the purchase. The information sources which they found most helpful are "Recommendation by friend/relatives" at 52%, closely follow by "consumer reports". Traditional media such as advertisement and magazine scored only 16% and 7% respectively. Reverting to the communication cues offered by video, social media could be a means of creating shareable content to drive discussion around a new product.

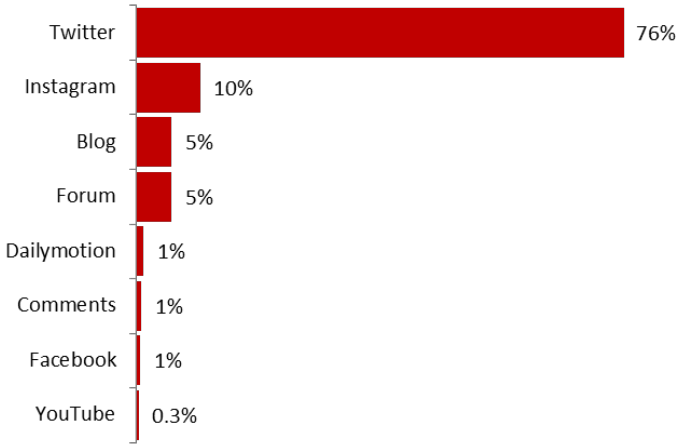
Figure 7. Research timelines for intended purchase



Who to target

It might be possible to deduce how, who and where to target the audience to possibly widen the reach to extend opportunities. The social listening findings showed that DIY related topics are mostly discussed in Twitter, Instagram, Blog and forums in the past 12 months. 79% gave neutral sentiment; 19% gave positive; and 2% gave negative. When we drill down to “mower” specifically, forums are the most popular media with 73% share.

Figure 8. Social media channels when searching DIY topics



Based on the gaps in social discussion alongside the survey and video response, Chervon were able to deduce two key tactics in their targeting across social for launch;

- Seek influencers / promoters / advocates to talk about Chervon products in wider circles of Social channels
- Engage environmentally conscious bloggers / socialites to talk more about Electric ride on mowers, potentially sharing from inspirational leaders i.e. Tesla to grow their own following and use these as opportunities of channels to expose Chervon products.

Broaden the category

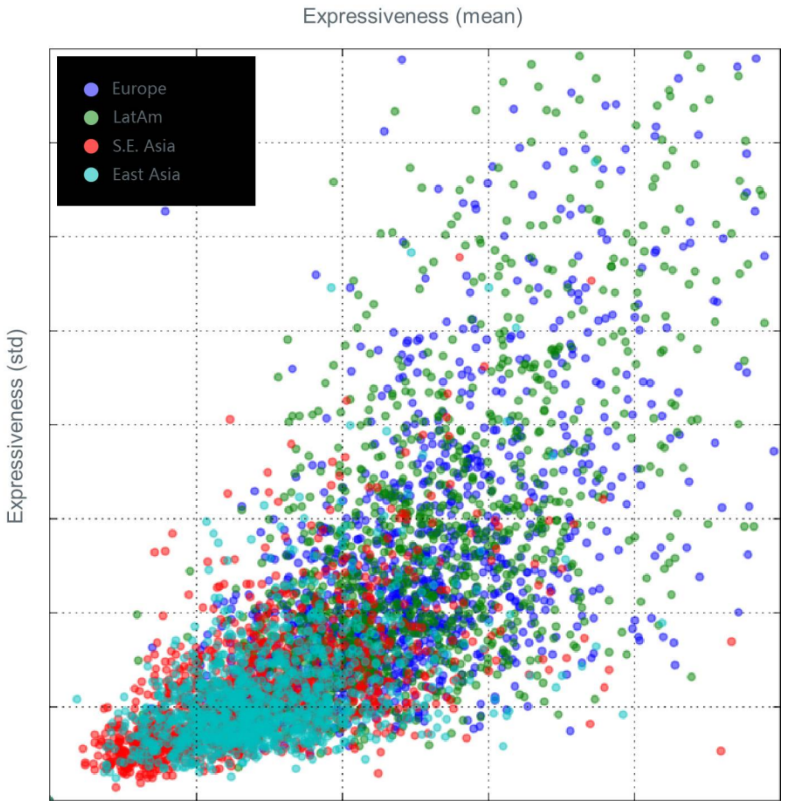
Off the basis of the social media component of this study, Chervon plan to search higher levels of social media keywords, e.g. DIY gardening, DIY power tools, DIY outdoor maintenance, etc.

They will also look at means of ROI to better understand where to best place promotional material, including word of mouth.

Facial coding: Building emotional understanding

Next we looked at the emotional responses of participants when they watched a 30 second advertisement about an existing Chervon push mower. Cultural context is imperative across research broadly, however with facial analysis this is vital. Cultural norms allow for proper contextual analysis and with over 4.8 million faces analysed (over 2 million of these in Asia) across 72 countries, this has been well developed.

Figure 9. Affectiva expressiveness regional comparison



By way of an example, pet care and baby care ads elicit the most enjoyment in the US, where as in Canada it's cereal ads.

What captivates the audience?

Within this study, expression, one of the key 10 metrics, rises to an initial peak in the advertisement when the inner components of the mower battery are shown and explained. If we come back to the earlier findings of the need to overcome concerns and myths about power source, this rings true. (See figure 10.)

However, when looking at the overall scores, the middle section is overall the weakest. This suggests that there is scope to cut this section down or focus on only few details to more positively engage viewers.

It then rises to the ultimate peak at the end, where the call to action is featured. However, against overall benchmarks this advert would benefit from a stronger call to action to heighten this reaction and instil recall and therefore proactivity in prospective buyers.

Figure 10. Expression across advertisement duration

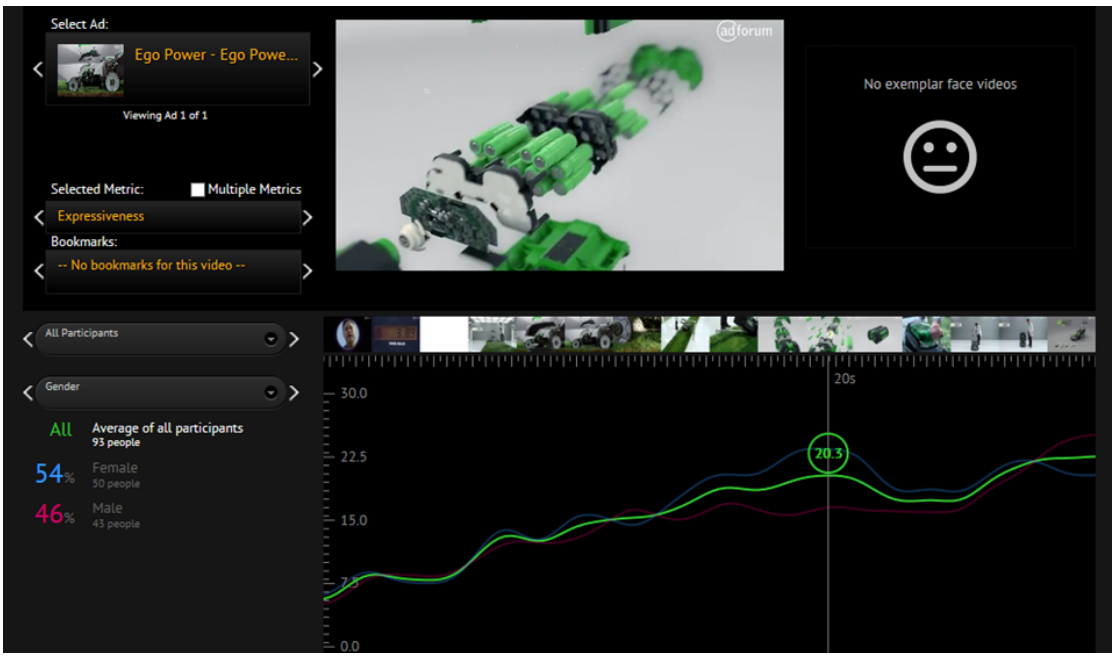
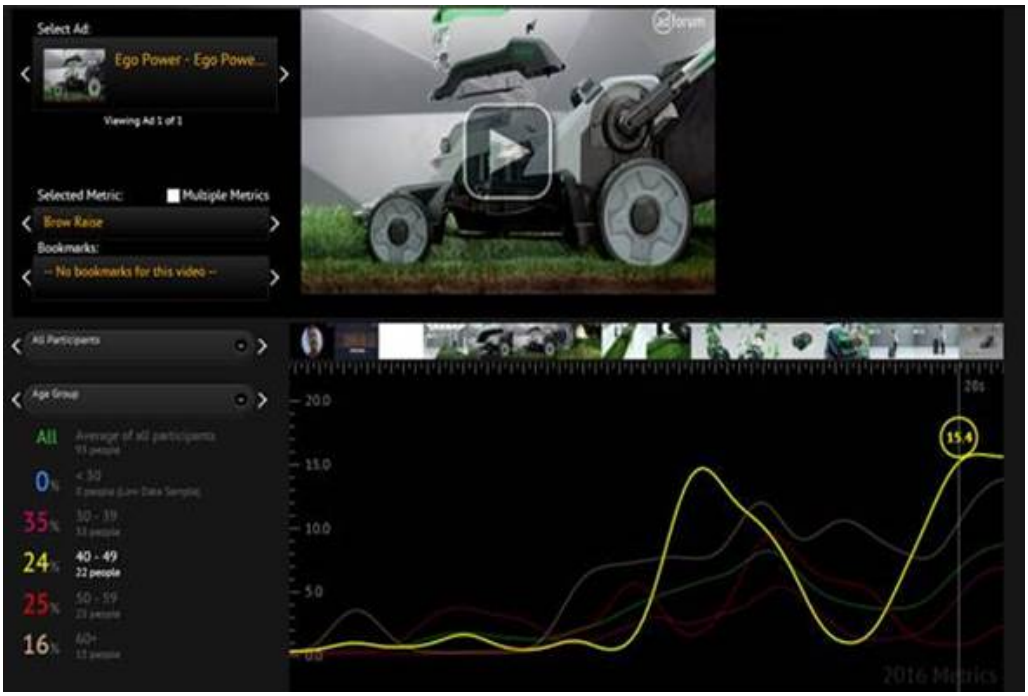


Figure 11. Overall scores for tested advertisement

When breaking this down by age bracket, 40-49 year olds had higher interest and were affected most by the advert.

Figure 12. Age group brow raise reaction overview



Running a deeper dive for this age group to understand their motivations is a route that the client is considering taking closer to launch.

Data appends: A passive portrait

And finally, to obtain further details of those who shown interested in battery powered mower found from our survey, we appended a further layer of data using Lightspeed's AmplifyR appends solution. Taking the IDs of those who told us in the survey that the ride on battery powered mower appealed we looked at the Experian Mosaic US coding that is overlaid on our panel. This resulted in a concentration in code 12, which is entitled as 'Blue Sky Boomers'. This gave us useful additional information of the type of audience through this category description.

According to Experian Mosaic, Blue Shy Boomers mainly refers to, older, empty-nesting couples and singles in anticipation of their retirement years, usually high-school educated. They are homeowners who tend to live in modestly-priced houses, and have settled in small towns, seaside beaches and lakefront communities. They can be found spending much of their leisure time outdoors, these households enjoy fishing, boating, hiking, hunting and gardening specifically. With childrearing days behind them, they have the discretionary cash to purchase things like boats, campers and pickup trucks, yet they can be price sensitive consumers when shopping generally. And for them traditional media still reigns supreme, but Blue Sky Boomers are becoming increasingly comfortable with the Internet.

Summary

By expanding the standard research methodology, one project made it possible to reach a clear direction for the client to proceed faster and more cost effectively than splitting into several projects. More so, all of this was achieved from China, 12,000km from the new market the client is focusing on. This was discovered as not only a new market but a hybrid of DIY and automotive for Chervon to tackle.

Key, actionable insights from the study were;

- The market is open to an electric ride on mower but myths and fear of the unknown need to be tackled head on pre and during launch to market.
- User content will play an important role in doing so, whereas mass marketing is not relevant given this niche market and how little social media or search content is currently in place.
- By adding video, the study also uncovered deep emotional drivers around why respondents feel they want ride on mowers, electric ones specifically and where product development could take the new machine.
- Both video and social then showed how to engage a broader audience by identifying who to work with and where to place the feeds to maximise the reach.
- Facial coding verified the interest points for communication to be based around and offered some clear direction on what to review when TV advertising is being prepared for launch.
- Although the product does not yet exist some of the key themes of environmental friendliness, being a status symbol and package upgrade/ personalisation allow Chervon to look at other existing products in different categories that share these features too, such as automotive. This will assist in the rollout in a new market in a totally new area.
- A qual/ quant hybrid now exists for researchers to transcend borders from a distance and access reliable data beyond a traditional survey approach, saving money and time.

This started with a basic business opportunity question from Chervon, a DIY tools manufacturer based in China – is there a market for electric ride on mower in North America? Within a fortnight, with little investment or time spent, we have discovered, confirmed a lot and provided focus, linking up the client with a group of potential end users and requirements. The traditional way would probably have been a market visit plus spending with consultancy firm plus qual / quant MR plus secondary research, with a lot of man-hours involved, be it using in-house resources and/or external agencies, even before the trail phase. This is particularly true for a new category that does not currently exist in the targeted markets.

As a research user, Chervon were able to see the impact new technologies can have not only on their own, niche requirements but on the research industry as a whole as well as answering to new consumer behavior norms. Integrating methodologies within market research practice enabled Chervon to understand consumer needs and behavior at a higher level than before. As a highly vertically integrated company growing from OEM to OBM in the global market, Chervon was anxious to explore local needs to create value for customers just about right. Integrating consumer data become more important when looking to enter a completely new market for. Any concerns about the multitude of data sources and methods blurring the lines were quickly overcome, as each component played a part in building on the view of the market, easily and efficiently. Working in this way helps brands to avoid or reduce the potential risks with trusted, broad data to base important decisions on at the very beginning of the whole value chain.

The Authors

Jeff Tsui is VP, Commercial, Lightspeed, China

Alex Tu is Market Insights Director, Commercial, Chervon, China

Big Data, A1, Machine Learning and Behavioural Science

The case of healthcare, energy efficiency and hospitality

Pascal Bourgeat • Burt Langton • Paul Cording

Introduction

Over the last three years, the deep learning version of machine learning has made more inroads into learning systems than all the artificial intelligence ever developed, particularly to understand image, voice and language processing, etc. However, a general intelligence system mirroring human intelligence must not only have our phenomenal learning abilities but also develop high level *social-cognitive abilities*. Machines can tell us what customers are looking at and feeling inside, second by second (eye tracking and facial coding), and doing (time and pathway in shopping centre, navigating store, supermarket alleys, airline lounges and online buying behaviour). Machines, however, cannot tell us directly what is going on with people at a deeper level, what cognitive and neural mechanisms are at play in what type of physical, cultural and social environment and what this means for what organisations want to achieve. Machines are still the dummies of the social-cognitive world.

Nevertheless, the defining characteristic and strength of 'supervised' machine learning systems is their ability to 'see' *patterns* that humans cannot see because they are hidden in massive amounts of data (with little signal and lots of noise). So despite their social-cognitive limitations, machines can see simple or complex patterns in behaviour and its antecedents that we cannot possibly see. But we can make sense of those patterns, especially if we have a razor sharp and systematic understanding of how behaviour works. Research companies don't have to wait for machine learning to learn unsupervised and the development of true general artificial intelligence to see and learn from those patterns. Current machine learning algorithms can predict well but the patterns they identify can become usefully descriptive (what is really going on here) and prescriptive (what does it mean for client) once a suitable behavioural lens is used.

Behavioural science has much to say about attitude and behaviour, how we form impressions, form preferences and make decisions/choice. The convergence of the systematic use of machine learning and behavioural science principles on all sorts of data sources (including surveys) can be a rich source of insights − insights that are available now and do not depend on how well machines and systems of the future capture and mimic our deepest and best cognitive abilities.

This paper presents a range of insights and their implications for commercial or public policy that are directly drawn from a process of confronting machine generated patterns (from classification trees as a frequent form of supervised machine learning and Bayesian networks) with behavioural science principles from cognitive and social psychology, decision neuroscience and bioenergetics/ neurophysiology. We list our five top tips for researchers (data scientists, consumer researchers, behavioural scientists, etc.) to get the best value for their private and public sector clients from the combined use of machine's eyes and human brain power.

Machine learning and patterns

This is a quick summary of the types of machine-generated patterns that have been used for this paper beside behavioural science which provides a body of knowledge from experimental research (psychology, ethology, neuroscience, economics) over many decades.

Machine learning algorithms

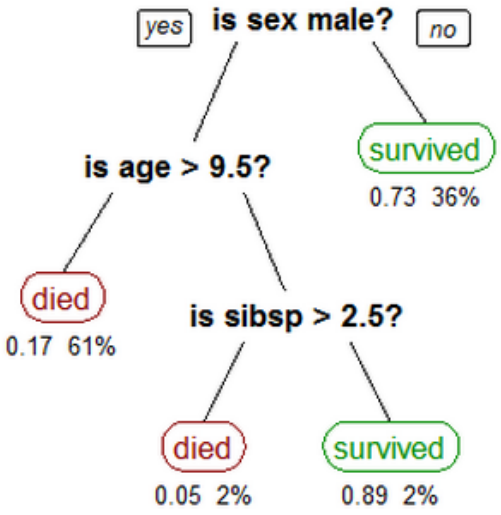
Classification and regression trees (CRT or CART) are one of the most used machine learning algorithms (MLAs) or systems. They are used by tax authorities, online retailers and technology companies, intelligence services, banks and financial institutions, online medical diagnosis and medical researchers,

political and campaign researchers and any organisation interested in predicting and scoring (tax evasion, national security risk, financial risk, voting or online searching behaviour, buying behaviour, mortality and morbidity likelihood, etc.)

CRT looks for patterns in data (scaled, ordinal or just nominal) to explain variation in target variable (e.g. behavioural). CRT creates sequences of questions that explain how individuals/cases end up in specific classes (the sequences of questions form the branches of the tree)

The following tree accounts for the survival of passengers aboard the Titanic and can be derived from the data describing survival and passenger description on the Titanic.[1)]

Figure 1. Survival tree of Titanic passengers



A tree showing survival of passengers on the Titanic ("sibsp" is the number of spouses or siblings aboard). The figures under the leaves show the probability of survival and the percentage of observations in the leaf.

CRT works as a supervised learning system whereby a set of data is used as a 'training set' so that patterns (the tree structure) leading to the target variable can be established. The patterns are stored in algorithms that can be subsequently used to score and predict individual behaviour or status or whatever the target variable reflects.[2)]

Scoring and prediction are often the most useful outcome of MLAs regardless of the content of the patterns they identify (what the trees look like as sequences of questions). Regardless of predictive ability, the content of the trees, however, is important because its interpretation can provide insight into behaviour or status (e.g. tax evasion, national security risk, financial risk, disease-related risk, buying behaviour, etc.). There are many types of MLAs who identify and capture patterns into algorithms in different ways and provide different results. Some MLAs may also be better at providing 'true' tree structure for real insight into behaviour whilst others may be better at prediction although their selection of predictors is biased.

Within a given MLA, parameter specification like the amount of pruning of the tree (avoiding over-fitting), its depth (the complexity of question sequences), the size of branches that can be further split into more branches or terminal nodes (the leaves of the tree), etc. There can be infinite alternative tree structures so there is always the question of what they really mean and what they are telling us about behaviour, real or spurious. In particular, a clear problem of MLAs is whether the classification trees are spurious because of over-fitting (i.e. at worst a very large tree would provide perfect classification for all cases with an almost equal number of leaves/nodes).

Baynesian networks

Like MLAs, Bayesian (belief) networks (BNs) seek to identify complex structures relating a set of variables to nominal or continuous outcomes (e.g. membership of a customer segment or frequency or intensity of a specific behaviour). BNs however are not about mere correlation. They rely on axiomatic rules of establishing cause and effect from, for example, thoughts and building blocks of mindset to overall impressions, disposition or behaviour. BNs analyse probability densities (distributions) of variables on the basis of conditional independence, ideally as manipulated densities (e.g. from randomised trials) rather than conditional densities (as directly observed in populations).[3)]

BNs are particularly useful to disentangle relationships that potential predictors have with each other as we try to establish their impact on an outcome variable (e.g. impression, preference, behaviour, etc.). For example, my thoughts about brands in a category shape my preferences for these brands but these thoughts are also related and can shape each other as part of the cognitive structure accessible from memory. BNs sort out those complex patterns and are very well suited to peek into cognitive structures elicited from surveys. As with CRTs, cognitive structures are provided 'as is' by BNs and questions should be raised about how real or spurious they are. Survey data is not perfect and tapping into the cognitive structure through a series of survey questions can potentially distort the very data that we try and capture. Yet, despite the fact that perfect psychometrics are elusive, BNs are able to identify causal structures (i.e. direction from predictor to predictor and from predictors to target outcome like attitude or behaviour).

In the following sections we present a range of application of the combined used of machine-generated patterns (from both MLAs like CRT and BNs) and behavioural science principles to evaluate the patterns and behavioural insights to further the commercial or social policy goals of clients. Three sectors are investigated: energy efficiency and electric vehicles in NZ, healthcare and patient experience in Australia and customer experience online and in hotels in various countries of South East Asia.

Application: The Energy Efficiency and Conservation Authority

The Energy Efficiency and Conservation Authority (EECA) is the government agency that works to improve the energy efficiency of New Zealand's homes and businesses and encourage the uptake of renewable energy.

EECA has been running a comprehensive survey research program to keep track of reported behaviours, attitude and reported advocacy, mindset, specific beliefs around energy efficiency and conservation, attitudes towards renewable energy (e.g. solar) and alternative transport fuels (e.g. electric vehicles).

Traditionally energy efficiency messaging has focused on reducing bills. A different route to encourage energy efficient measures pursued by various countries, including EECA, has been to frame energy efficiency as part of our desire for a warmer and more comfortable home. EECA have also focused on providing balanced information to the market through social marketing to empower people with knowledge to make informed energy efficient choices.

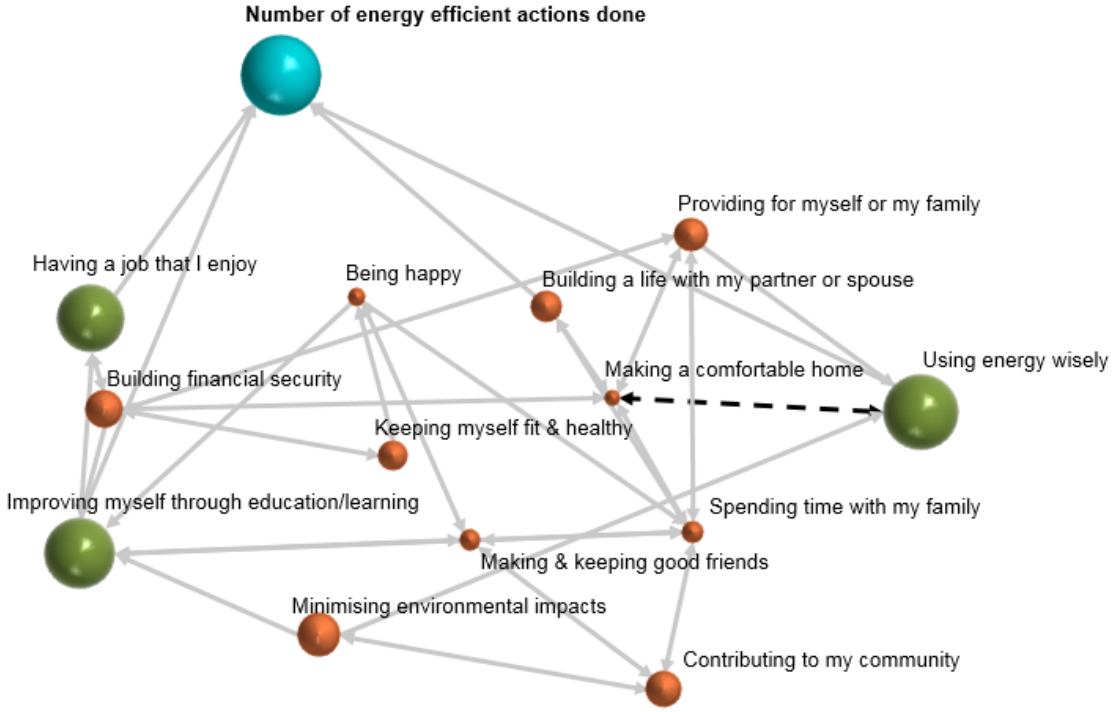
We wanted to understand what was related to efficient energy use within the broader mindset of people's life priorities. Over several years, survey respondents have been asked to indicate core priorities, lesser priorities and no priority in their lives. The Bayesian network (BN) illustrated in figure 2 shows how different priorities related to each other to create mindset. The BN also identifies which of those priorities causes or contributes to the number of energy efficient measures that people say they have implemented (e.g. installing home insulation, using energy efficient light bulbs, improving hot water, tyres or solar panels, purchasing an electric vehicle, etc.).

The network reveals that as a mindset, using energy wisely is clearly related to 'minimising environmental impact', but the link to 'making a comfortable home' is not readily present. The latter however is expectedly related to 'building financial security' and 'providing for self and family'. If we force the network to reveal more tenuous links (the dotted line) within the mindset of the public, 'making a comfortable home' is related

to 'using energy wisely'. This shows that the hedonic 'making home comfortable' through 'using energy wisely' is not a dominant feature of the current mindset (i.e. only present in the mindset of a smaller proportion of the population) whereas the pro-social minimising environmental impacts is. This can also be related to the fact that some aspects of using energy wisely e.g. insulation are a one-off action, rather than an ongoing activity.

The contribution of using energy wisely (as a priority) to the number of measures people have adopted is evidence of a link between motivation and behaviour. However, the network also reveals two aspects of people's mindset of life priorities that equally impact behaviour around energy efficiency, above and beyond motivation: having a job that they enjoy and improving themselves through education/learning. Both of these reflect underlying barriers (or enablers) to adopting energy efficient measures: the former through a sense of optimism (and possibly choice/affordability if low earnings correlate with poor job satisfaction) and the latter through awareness of the issues around energy use and knowledge of possibilities and choices available, which is an area that EECA has focused on through its marketing efforts.

Figure 2. Bayesian network of lifestyle priorities and their impact on the number of energy efficient measures adopted



From a behavioural perspective, the network clearly reveals that motivation and ability are both equally important for behaviour. Campaigns and interventions often focus on motivation, assuming that motivation offers more scope to leverage behaviour than ability. The current weaker link between 'making a comfortable home' and 'using energy wisely' confirms that 'making a comfortable home' is only one motivation and different households will respond to different motivations. It also confirms that other aspects of people's mindset about life act as strong enablers of energy efficient behaviour for some people and strong barriers resulting in no or modest energy efficient behaviour for others. EECA's focus is on educating the market, providing subsidies (where possible) and motivating the market through illustrating co-benefits of energy efficiency such as monetary savings but also having a warm and comfortable home, less sickness at home and having energy for all in NZ in the future. Driving motivation and enabling desired behaviours (pulling down barriers, creating opportunities) is an appropriate strategy for behaviour change.

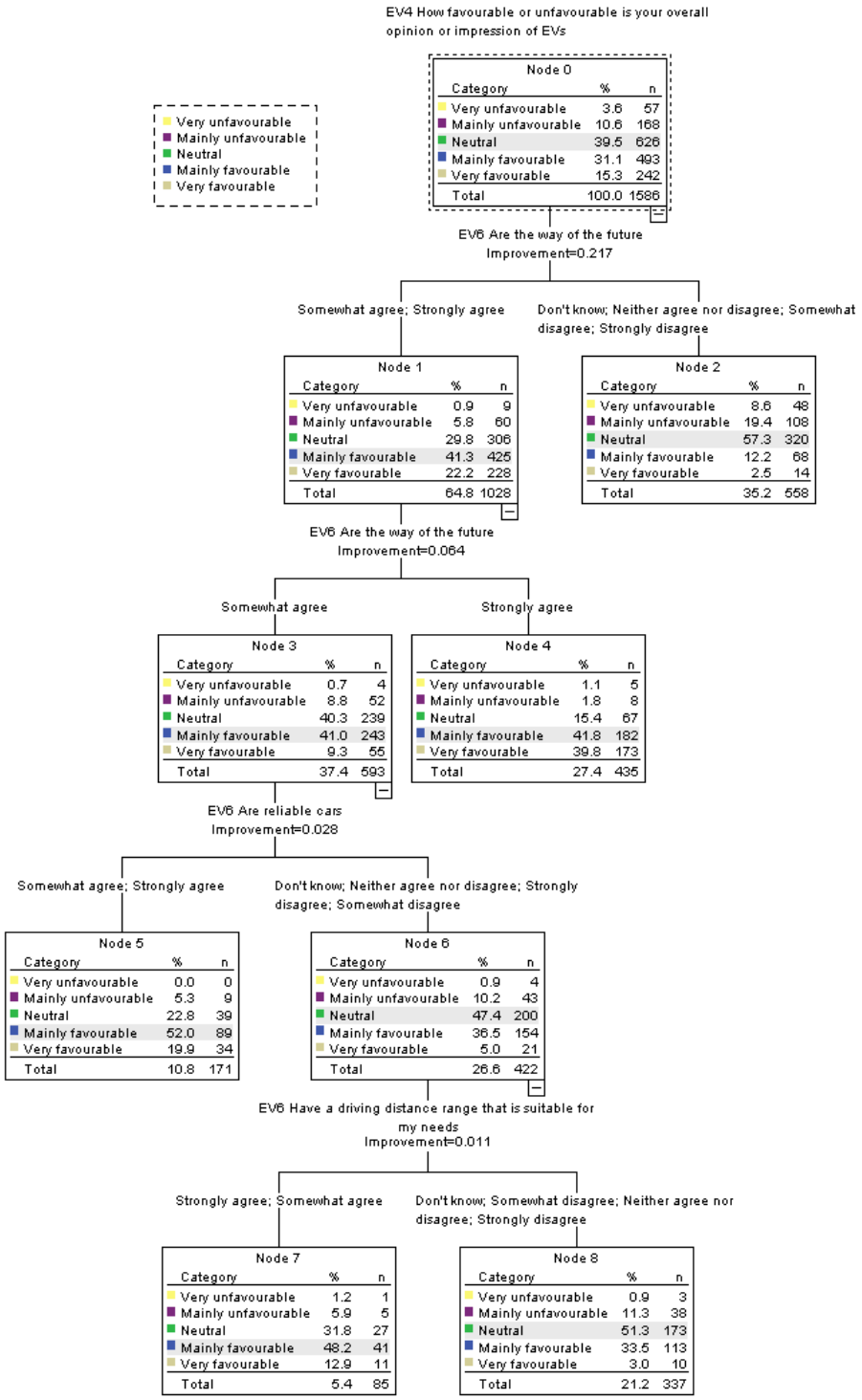
Attitude towards electric vehicles

For a limited period of time, the same survey also tapped into the knowledge and attitude of electric vehicles (EVs) among the public. We identified mindset and its link to favouring electric vehicles using a

MLA (SPSS CRT). Survey respondents provided an overall favourability rating towards electric vehicles and their beliefs around various positive and negative aspects of electric vehicles were also measured (e.g. have lower maintenance costs, driving range, time to charge, preserve the environment for future generations, affordable price, etc.). One strategy used for the diffusion of technological innovation is to highlight its benefits (and minimise its perceived downsides) vs other options (here petrol cars).

The expectation about what the MLA's tree would reveal was that it would highlight which benefits and negatives most impact favourability towards EVs. The tree structure in figure 3 shows this is indeed the case but with a clear twist. The primary driver of favourability is not a specific perceived advantage or disadvantage of EVs but their feeling that EVs are the way of the future.

Figure 3. Classification tree of favourability to EVs as a function of beliefs about EVs



Behavioural science provides a clear framework to make sense of this both in terms of the shape of people's thinking, feeling and intentions towards EVs as well as practical implications to further develop the diffusion of EVs in the population.

The *affect heuristic (or shortcut)*[5)] is a cognitive mechanism by which people engage an affective evaluation of the situation *before* accessing and processing thoughts/beliefs directly related. In other words, our evaluation of EVs is the result of both what we feel about it and what we think about it.

A number of specific pros and cons shaped the public's disposition towards EVs (e.g. driving distance range without recharging battery, time to recharge, easily available recharging points, positive impact on environment, reliability, etc.). However, it also identified the presence of a direct and immediate 'affect shortcut' in 'EVs are the way of the future'. This effect impacts disposition above and beyond that of specific positives and negatives of EVs. As would be expected from such an affect shortcut, perceived benefits are enhanced and perceived disadvantages minimised among people associating EVs with the way of the future (and vice versa for other people).

The implications from the affect shortcut in the case of EVs are that the road to gaining favour with drivers is not simply about promoting the benefits and avoiding the negatives. Experienced or reported negatives like insufficient autonomy or poor availability of recharging points are likely to inhibit purchase behaviour among the public anyway. However, initial perceptions (of positives and negatives) and willingness to engage are coloured by people's view of EVs as the way of the future.

Importantly, the way of the future is implicitly linked to what the behavioural norm (choice of EVs) will be at some point in the future. Visible cues of the increasing use of EVs, information on their penetration of the automotive market and messages designed to reinforce their overall image as the way of the future all contribute to faster take-up of EVs and support the effective communication about perceived benefits of EVs vs other types of vehicles.

The dynamic or self-feeding *leverage of the social force* is that the more people there are who perceive (from their environment) that EVs are the way of the future, the more likely it is that everyone else will see it the same way. This is the way the attitudinal and behavioural norm is shaped. In addition, the **leverage of the affect shortcut** is that the more people see EVs as the way of the future, the more their perceptions of positives and negatives are coloured by affect: the more believable and clearer the benefits become and the less impactful the negatives become.

Application: Patient experience in Victoria

The Department of Health in Victoria conducts patient surveys to evaluate the experiences and overall impression of emergency department (ED) and planned admission patients whilst in hospital. The fifteen minute surveys are available in fifteen languages beside English. They are administered by mail-out or online two to six weeks after patients leave hospital.

Patients provide feedback on anything from admission, the food provided, their experience of the ward, the doctors, the nurses, other health professionals, multiple aspects of care or the tests and procedures to the time patients are discharged with some room at the end of the survey for feedback in the patients' own words.

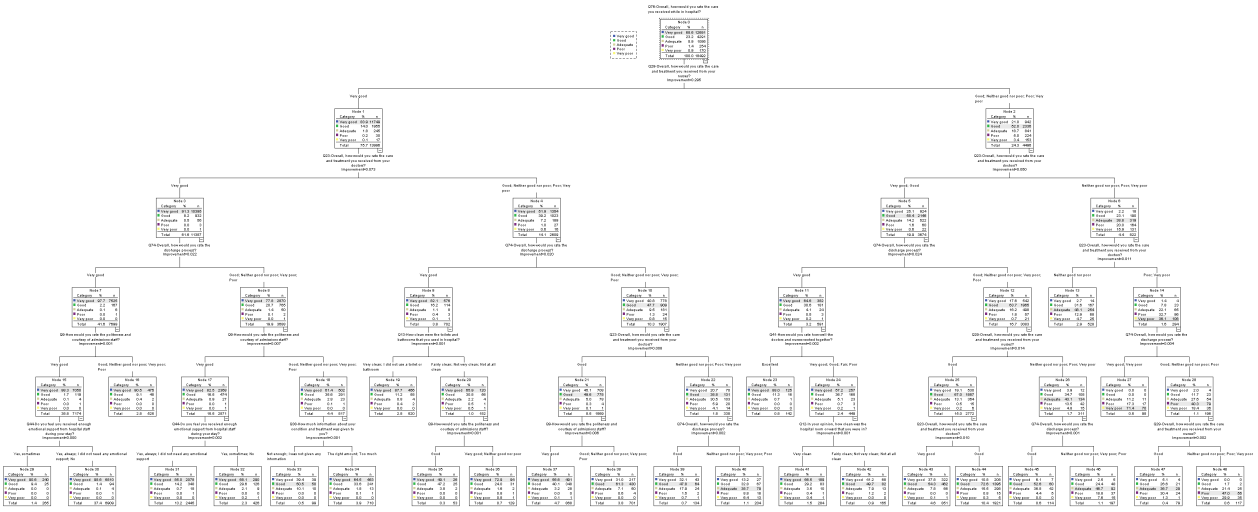
Using MLAs we examined planned admission vs ED patients together and separately to find what the pattern of their experiences looked like with regards to those experiences that shaped their subsequent impression of their hospital stay.

Behavioural scientists raised this very question about how our impressions (from the memories of our experiences) are shaped. They found out from experiments that lasting impressions are **not** shaped as the average of our experiences (e.g. positives offsetting negatives) but can be subject to a rule called 'peak-

end'. That is, our overall impression is shaped by what happens at the start (but generally not much), really positive or negative experience sticking out as a peak – or a trough (much more) and the experience at the very end (much more); hence the name peak end.[6)] Peak-end is *not* a universal rule to explain the shape of all our memories of sequences of experiences but it provides a cue about what we can expect to find out in the patterns that MLAs put before our eyes as they sift through a large set of data points (here about four million data points in the case of the patient surveys for adult in-patients alone).

The tree in figure 4 provides an overview of the different patterns found among the combined sample of planned admission and ED adult in-patients.

Figure 4. Classification tree of patient overall satisfaction as a function of hospital stay experiences



There are many observations to derive from this tree but for the purpose of this paper we will focus on the MLA's initial split of patients' overall impression of their hospital stay using their experiences of nurses and then those of doctors. Particular positive or negative experiences that stick out in memory don't appear to shape patients' subsequent impressions possibly because there is a constant or recurring pattern to the presence and work of nurses and doctors around patients. Patients in hospital are also subject to feelings of uncertainty (their condition, their outcome, their future) and vulnerability (they have very little control over their life in hospital) and this may account for the large impact of emotion colouring their memories rather than peaks of positives or negatives.

For some pockets of patients, however, particular experiences do shape their impressions quite strongly but this remains marginal because relatively few patients are affected. Cleanliness of toilets and bathrooms or cleanliness of hospital room or ward are two such experiences and they may well stick out in patients' memories because of the likely effect of cognitive dissonance between visible dirt/uncleanness and hospital/health.

The tree however detects the effect of both the start and end of hospital stay on patients' overall impression. Figure 5 shows that start (experience with admission) and end (discharge) don't add much to hospital experience beyond nurses and doctors when all experiences are rated 'very good'. However, start and end have significant negative influence on overall impression when experience is not 'very good' (i.e. from 'very poor' to 'good').

Figure 5. Impact of peak-end experiences on overall patient satisfaction

Experience	Overall impression index*	
Care and treatment from nurses: very good	122	and
Care and treatment from Drs: very good	133	and
Discharge process: very good	143	and
Politeness and courtesy of admission staff: very good	143	

Experience	Overall impression index*	
Care and treatment from nurses: very good	122	and
Care and treatment from Drs: very good	133	and
Discharge process: good ⇔ very poor	114	and
Politeness and courtesy of admission staff: good ⇔ very poor	89	

** index based on % 'very good' experience overall with hospital stay for total sample*

A key question that arises from the pattern identified by the MLA is why discharge is not more impactful than admission? As is often the case, the answer lies in the fundamental principle of behavioural analysis that behaviour is as much about the person as it is about the context. Behavioural mechanisms like peak-end are modulated by the context which interacts with other facets of the person's cognitive and emotional processes: admission can be a lengthy process (waiting for ED) and admission is also a time when people can be uncertain, vulnerable and in pain. The context validates why both start and end appear in the tree. The role of uncertainty and vulnerability is further evidenced by the impact of two specific factors appearing in the trees for ED patients (i.e. patients for whom the whole experience is unplanned and mostly unexpected):

- Whether patients feel they have received enough emotional support from hospital staff during their stay.
- How much information they receive about their condition and treatment.

From this perspective, the start of the hospital stay for ED patients does appear like a sticking point in memory because it is the time where emotional support and information tackle high level of uncertainty and vulnerability through our emotions, feelings and thoughts.

There are very good reasons as well as to why discharge would stick out in memory apart from its position at the very end of a sequence of experiences. The discharge may be delayed because the doctor has not had the opportunity to speak to family or carers or transport is unavailable or the patient's home situation has been addressed, e.g. meals or mobility aid, etc.). All of these have the potential to create significant negative emotion for patients and stick out in memory.

We have only examined a section of the patterns presented here about patient experience (to protect the confidentiality of the study) but from this extract there are some clear policy implications:

1. Patients may be in hospital for a few days but what happens in the very last few hours (and to a lesser extent the first few hours) can have a disproportionate negative impact on their impression of their hospital stay (relative to total time in hospital). This also applied to patients who rate their experience of nurses and doctors throughout their stay as very good.

2. Hospitals see care in the joined work of doctors and nurses as the heart of what they do. However, from the point of view of patients, what happens at the start and at the end, right before patients are 'on their way', does colour their overall impression disproportionately at times relative to their time in hospital. In this case government intent on improving services to patients can more closely align with the perspective of patients.
3. The patient's context bears significantly on their lasting impressions because some of the factors shaping the impression of, for example, ED patients will be different from the factors of planned admissions, pregnant patients or parents whose child is the patient. Nevertheless, some of the factors shaping ED patients can be expected to be the same across all types of patients. Big data being big enough there are general and specific behavioural insights to work out policy at the right level of granularity for each type of patients. On the other side, small data means effects can be averaged across patient types and remain hidden.

Application: Customer experience online and in hotels

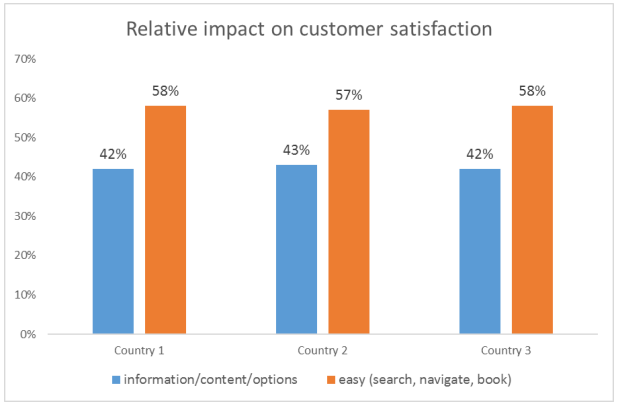
Experience of online purchasing

Travel services (airline tickets, tours and hotel bookings) were the most commonly purchased items online in South East Asia (SEA) according to a 2014 report.[7)] A 2017 Ipsos survey of online retailing across SEA, North and East Asia and Australia confirms the top position of airlines and hotels for online purchases but now followed closely by finance and fast fashion (in Asia). Regardless of sector, experience through booking or purchase process impacts customers' impressions, preferences and future behaviour beside the actual experience of the service or product purchased online. Online processes are designed to convey the offers developed by marketing teams in tune with customers' needs, wants, price and value points for products and services. These are the components which drive customers' *motivation* to search online, pay attention and notice, engage and stick with an online process. However, one key principle from behavioural science across multiple disciplines (from cognitive psychology to ethology and neuroscience) is that motivation is offset by anything that makes the decision making (or processing the situation generally) difficult and/or creates a *negative affective state* (emotion, feeling or even mood). *Lack of ease* in online behaviour can easily inhibit behaviour (customers abandon search, comparison, sequential process, filling in details, etc.) not because they are unmotivated but because (lack of) ease and negative emotion means customers require stronger motivation. The more effective road (but often not the preferred one) to deliver better customer experience, overall satisfaction (and sales and future repurchase) is to prioritise ultra-simplicity and ease throughout all processes and avoiding negative affect rather than tasking the marketing team with creating more motivating offers or content.

Ipsos conducted a series of surveys in SEA looking at customer experience and satisfaction with an online booking process. Using the SPSS CRT machine learning algorithm, we investigated the differences between countries as to what aspects of the online process most impact people's overall satisfaction with their experience. In addition to a detailed tree structure showing the patterns of experiences and satisfaction levels, the MLA also provides a normalised score of importance for all the predictors of satisfaction (with the online process). CRT showed that *in all countries* perceptions related to the ease of the process were significantly more important than perceptions related to the information, content or options available to drive customer satisfaction.

The comparison is presented in figure 6 for three of the countries surveyed (the identity of the countries has been removed to protect confidentiality).

Figure 6. Relative impact of content / information vs. process on satisfaction with online booking across SEA countries



Customer actual experience in the same survey revealed that 40% to 50% of them rated their online experience on statements related to ease of use (search on site, navigate, book, etc.) as 7 or less (on a 1 to 10 scale where 1 is strongly disagree and 10 is strongly agree).

As researchers sharing our observations and analysis with clients, we have observed the same issue with ease in multiple sectors related to online retail (and other sectors like banking, finance, telecommunication, traditional retail, etc.) and also the same leverage on changing customer experience and behaviour. This is not alien to our own experiences as shoppers, customers and service users where repeatedly we encounter design that is not intuitive, lay-out that requires more cognitive processing than it should, sequences that are not communicated, information requested at an inopportune time or designed in a straightjacket.

The substantially strong leverage offered by perception of ease (vs all other perceptions that are intrinsic to customer motivation and online behaviour) and the room for improvement in ease suggest that the client could usefully look at CX design in priority to marketing content to develop their customer base and sales.

In the quest for consumer or customer-centricity, design thinking (and more generally systematic behavioural analysis and review) provides more scope to effect experience and behaviour change than what marketing can do (e.g. inserting more online promotional offers). Organisations can also drive up the appeal of their content/offers. But these efforts often fail to get over the wall of diminishing returns unless organisations initiate some form of disruption.

CX and its Design Thinking toolbox however need more than a large dose of empathy with customers and observation and interviewing skills. An effective behavioural lens is required to systematically identify the personal and contextual forces that constrain, inhibit or facilitate and encourage behaviour. In addition, CX design can benefit from the exploration of patterns identified by MLAs and interpreted in light of behavioural science principles.

Customer hotel experience

In surveys of customers' experience and satisfaction with hotels in SEA, we investigated which customer experiences led to a positive impression. A whole range of experiences from the time customers check in to the time they check-out were measured as well as satisfaction with the overall experience and all experiences across all hotels sifted through a MLA.

Because of peak-end rule we would expect to see the emergence of check-out as a significant factor in the overall impression of hotel customers in each country. Check-in and check-out together consistently contribute between one fifth to a quarter of total satisfaction. We would also expect to see check-out contribute more than check-in to overall satisfaction but the symmetry may simply reflect the inability of the

MLA to ascribe more value to check-out because of lower data variation in check-out satisfaction (around 75% of customers rate check out as 8-10 and about 95% 7-10). In turn stronger satisfaction with check-out reflects hotels' focus on check-out as a critical moment of customers' hotel experience as they make sure customers do not finish on a negative note. Given the impact of positive end experiences on subsequent memories and evaluation, a key question for hotels is whether they can create something positive (a peak) in check-out experiences rather than just creating one which is negative-free. This is in contrast with patients' 'check-out' experience where the complexity of the situation at the discharge point can more easily give rise to negative experiences. The focus for hospitals is to design experience so that patients' experience doesn't end up on a negative 'trough'. The focus of hotels is to design experience so that customers' experience ends up on a peak.

Figure 7. Relative impact of peak-end on hotel stay across SEA countries

Experience	Country 1	Country 2	Country 3
Check-in	11%	12%	14%
Check-out	11%	10%	10%
Other hotel experiences*	78%	78%	76%
Total	**100%**	**100%**	**100%**

** room cleanliness, hotel amenities, dining, etc.*

Conclusions and top tips for social and consumer researchers

AI is here. It is already everywhere but it is about to become much more visible in day-to-day machine-human interactions thanks to cognitive advances in robotics. Research people however routinely work to uncover, make sense, anticipate and shape our needs, desires, intentions, thoughts, emotions and behaviour. To achieve this demands that we develop and use the best of our abilities: our social-cognitive functions. These are not the types of functions that machines are likely to master, certainly not soon, maybe never. Some say hopefully never. However, machines are formidably equipped with AI systems, through supervised learning, to detect patterns which can be embedded into systems and used for many predictive purposes (medical and healthcare, aged care, education, finance, national security, tax evasion, retail and online retail behaviour and of course business and social policy research). When combined with our knowledge of behaviour those patterns can also provide rich insights into the mindset and behaviour of shoppers, customers, service users and citizens and help organisations achieve their commercial and policy objectives.

Here are five top tips to make the best of the combined use of machine learning AI and behavioural science principles:

1. MLA can be used on a variety of data sources including surveys as long as the data is reasonably big (usually 000s but preferably 000 000s or millions of cases). They are easy to run and can easily accommodate most kinds of data types as they come.

➔ Develop the habit to let machines find patterns that we can't see on all data available (single sources or mixed from surveys, client databases, open-sourced databases, etc.).

2. All MLAs are sensitive to parameter specification and changes can show different patterns. Make sure you understand how the different parameters do and how they affect tree structure and the strengths and weaknesses of different types of MLAs.

➔ If you have access to different types of MLAs, compare the patterns they provide. Key behavioural learnings are found in the tree patterns, not in the easy but information-poor ranking of all drivers.

3. Using behavioural science to change behaviour is always about understanding how the personal interacts with the contextual: how do our motivation and abilities shape our behaviour in interaction with our physical, temporal and social environment. Many cognitive mechanisms and functions (e.g. attention, memory, emotion, cognitive illusions and decision shortcuts, etc.) greatly impact our impressions, judgment and decision making. The patterns uncovered by machines often reveal

those same mechanisms at play as well as the opportunities and constraints they create for influencing and changing behaviour.

➔ Learn what those mechanisms and functions are. Unless we know them, we are unlikely to see them when we need to see them.

4. Behavioural science principles, especially the range of cognitive mechanisms at work in decision making help validate the patterns provided by MLAs and create real insights. However personal forces always interact with contextual forces so we need to keep in mind that an expected effect may not be taking place because of the particular place, moment and social environment.

➔ Always think person + context and don't treat behavioural science principles like dogma. Context creates 'exceptions' to behavioural principles that are all too often accepted as universal.

5. Many people around the world are busy trying to embed high level social-cognitive abilities into machines (robots and mainly AI running everywhere in the background of our lives). Research people depend on those social-cognitive functions to make sense of the patterns that machines 'see' but these functions are still a long way to be embedded in machines like they are in experienced humans.

➔ AI engineers need to cooperate with researchers and behavioural scientists to 'can' the type of cognition that research taps into when they create insight from patterns or configurations in data.

But in the mean time we'll continue working with whatever machine and human abilities we can muster.

Endnotes

1. The Titanic data are available at https://inclass.kaggle.com/c/titanic-survival/data.
2. Details about bagging and boosting, the key steps in identifying and testing branch structures in CRT can be found at http://www.stat.berkeley.edu/~breiman/wald2002-1.pdf and http://www.machine-learning.martinsewell.com/ensembles/bagging/Breiman1996.pdf .
3. Specific description of the directed acyclic graph (DAG) and the vortices made from the densities of all variables from which causality can be formally established are described in http://jmlr.org/papers/volume11/spirtes10a/spirtes10a.pdf
4. Details about Ipsos approach to establishing causal inference using structural search can be found in the appendix section of this link:
   http://www.ipsos-na.com/dl/pdf/knowledge-ideas/marketing/IpsosMarketing_Drivers_Analysis_POV.pdf
5. For a more detailed understanding of affect in decision making and the specific mechanism of the affect heuristic, see
   Zajonc R. B. (1980) Feeling and thinking: preferences need no inferences, *American Psychologist*, 35(2), 151-175.
   http://www.ideal.forestry.ubc.ca/frst524/06_zajonc.pdf and
   Zajonc R.B. (1984) On the Primacy of Affect R. B. Zajonc, *American Psychologist*, 39(2), 117-123
   http://www.ibl.liu.se/student/kognitionsvetenskap/729g02/filarkiv/ht11/1.297568/zajoncprimacy.pdf and
   Slovic P., Finucane M., Peters E., MacGregor D.G. (2002) Rational actors or rational fools: implications of the affect heuristic for behavioral economics, *Journal of Socio-Economics*, 31, 329–342
   http://faculty.psy.ohiostate.edu/peters/lab/pubs/publications/2002_Slovic_Finucane_etal._Rational_actors_orational_fools.pdf
6. Some of the original research supporting peak-end as well as a related effect (duration neglect) can be found at
   http://brainimaging.waisman.wisc.edu/~perlman/0903-EmoPaper/KahnemanFredricksonSchreiberRedelmeier_1993_WhenMorePainIsPreferredToLess.pdf and
   https://www.amherst.edu/media/view/231465/original/colonoscopy%2B2.pdf and
   http://www.cmu.edu/dietrich/sds/docs/loewenstein/WhenDurationMatters.pdf and
   http://web.mit.edu/ariely/www/MIT/Papers/duration1c.pdf
7. http://www.nielsen.com/apac/en/insights/news/2014/southeast-asian-consumers-flock-online-to-purchase-products-and-services.html

The Authors

Pascal Bourgeat is Director of Behavioural Science, Ipsos Australia.

Bart Langton is Research Director, Head of Marketing Quant, Ipsos New Zealand.

Paul Cording is Marketing Manager, Energy Efficiency and Conservation Authority (EECA), New Zealand.