

# COCA-COLA'S AGILE RESEARCH ADVENTURE

## IMPLEMENTING AN AGILE RESEARCH PROGRAM IN ASIA

Ron Gailey • Stephen Phillips • Pete Cape

### INTRODUCTION

This paper tells the story of Coca-Cola's partnership with ZappiStore and SSI and their successes and failures developing research solutions in Asia that we call "Agile Research" - research that is done in the spirit of fail fast and learn quickly, while providing results we need in hours or days, not weeks.

ZappiStore is an online research provider that offers an automated self-service platform for buying research services. SSI is a global provider of data solutions and technology for consumer and business-to-business survey research.

Fast and flexible tools have been available in the Western World for a few years, but not in Asia. The implication of this effort, assuming it works as we envision, is that The Coca-Cola Company and other marketers will have the ability to use quantitative consumer insights earlier and more often in the decision-making process in Asia. In forming this collaborative partnership Coca-Cola explored real-world applications of agile and cost-effective research, shifting the purpose of insights from validation to continual learning.

The collaboration was undertaken to explore a new set of problem-solving approaches. ZappiStore would provide tools that offer speed and cost savings in the early learning stages. SSI would provide sample and analysis to validate the reliability and limits of their online sample. Coca-Cola would provide guidance and content to be tested. All three companies would work together in an iterative process and collaborate for improvement and learning along the way. The promise of this iterative approach is profound as it should change how The Coca-Cola Company adapts to minimize risk through quick learning—giving the marketing teams the ability to make course corrections faster and check ideas in almost real time.

The full results of these efforts are shown below. Our goal with this paper is to share all learnings, successes and failures, across key countries in Asia so that the industry may benefit from our experience.

### BACKGROUND

The Asian markets continue to present global brands with both new opportunities and ongoing challenges. Indeed, this dichotomy is well understood by those involved in market research.

The Coca-Cola Company produces a significant amount of advertising in Southeast Asia. They have solid and reliable processes for ad development. However, the changing world demands better, faster and more cost effective work. In 2014 Coca-Cola sought to add speed and flexibility to their Ad Development and Testing toolset by forming collaborative partnerships with several consumer insights companies. The end goal was to develop an improved research process that builds on the current qual/quant tools with fast and cost effective quantitative studies which enable the marketing teams to more rapidly iterate in the early stages of new communication development.

If effectively harnessed, we believe that iterative research can offer the potential to shift how organizations think about the purpose and timing of quant and qual research in the development process; moving new quantitative methods earlier in the validation process as an important engine for ideation, creation and growth.

### THE AGILE NEED

Coca-Cola's research team recognized a potential opportunity to move new ideas from conception into development quicker by baking rapidly deployed quantitative consumer insights with iterative testing into the development process.

Existing, traditional research approaches offer a variety of innovations in terms of methodology, but lack the ability to deliver at the pace and price that business now requires. As stated by Ron Gailey, Director of Consumer & Commercial Insights at Coca-Cola Pacific, “*You can fail, but you need to fail early, fast and cost effectively.*”

Focus Groups are very helpful and provide a wealth of understanding, but using them in the Southeast Asia region to select between promising directions can be risky, as focus groups are already very small sample sizes and cultural tendencies towards agreeing with others in these settings (group think) may create even further problems. This can lead to frustrating resets when later quantitative work reveals issues. An agile and cost-effective quantitative approach was needed, but did not exist in Southeast Asia.

The online research provider has a strong suite of high quality and fast automated research tools that work well in many Western countries. It appeared to be a logical solution to the Coca-Cola need. Yet, there were significant concerns and barriers for successful deployment and adoption.

- Many researchers like the idea of change, but it creates angst. There is inertia to do things the same way. In short, the research community can be slow to adopt great and promising ideas because of the unknown. The Coca-Cola Company marketing teams and agency partners had similar concerns and did not want to invest in research that might undermine their success even if it was fast and low cost.
- The agile idea is very promising. Yet, it was not ready to be deployed in Southeast Asia, and there were challenges to get it functioning in optimal, automated form. Further, how consumers would respond to the tool in Southeast Asia was not known.
- The tools rely on solid panels of online consumers. Even if the idea was great, we are dependent on the quality provided by the agency providing the sample. Southeast Asia tends to have young and evolving online panels which could dramatically influence our findings. A great test that is executed on consumers that do not reflect reality could undermine the whole process.

This story is the result of three companies with three different needs who came together to help each other with a win-win-win. ZappiStores has products that had the potential to be a game changer in Asia. Coca-Cola needed agile research tools as the business moved faster and faster, but needed to be sure it would work reliably with consistency across markets. SSI has been running online panels for many years in Asia. Yet, they wanted to build confidence in their panels; to prove and validate them for reliability against other methods used, and also to learn how to improve their panels. Leaders from Coca-Cola, ZappiStore and SSI agreed to work collaboratively in an open and transparent way to explore and learn – with the aim of failing fast and improving quickly, and living by the very approach The Coca-Cola Company hoped to adopt. This story highlights the journey these three companies took to help one another shape industry change and benefit for their firms.

## THE AGILE TOOLS

The focus for this paper was the fast and efficient Ad Testing tools available on the online research provider’s site.

What the system enables is a fully automated approach to market research with a simple three-step process to set up any study in three to five minutes:

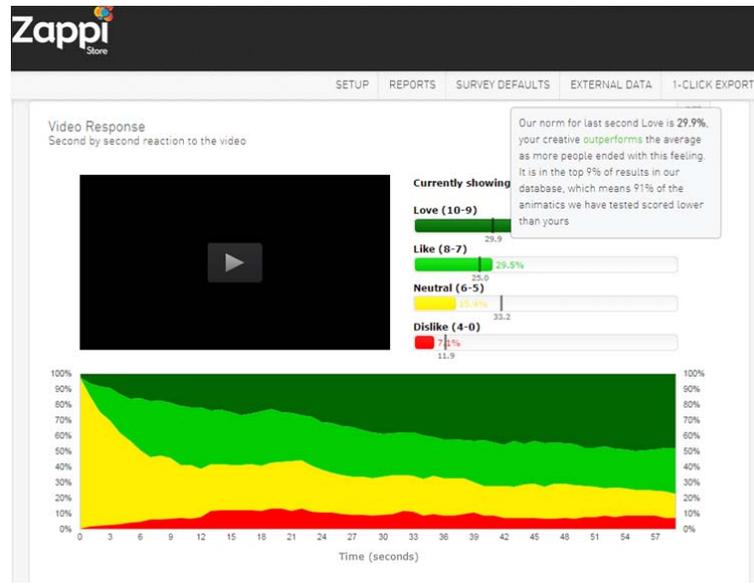
- Choose a target audience from a predefined list of demographic and behavioural audiences
- Upload creative (animatics, TVCs, etc.)
- Define appropriate survey parameters, such as messaging or other issues to be checked

The system automatically does the following:

- Generates a standard survey with amends as specified by the user
- Sends a survey link to relevant panellists using an API link with the panel partner
- Cuts off at the pre-defined sample size and checks on quotas and data quality
- Ports the data to a user-friendly dashboard
- Analyses of open-ended text and commentary on the results created using machine learning. (Example screens are shown below.)

In figure 1 we can see an image of an online dashboard providing second by second reactions to an advert, allowing clients to see what parts of the ad create the strongest feelings. Findings can also be compared to benchmarks from the overall database of ads tested or to a client database.

FIGURE 1.



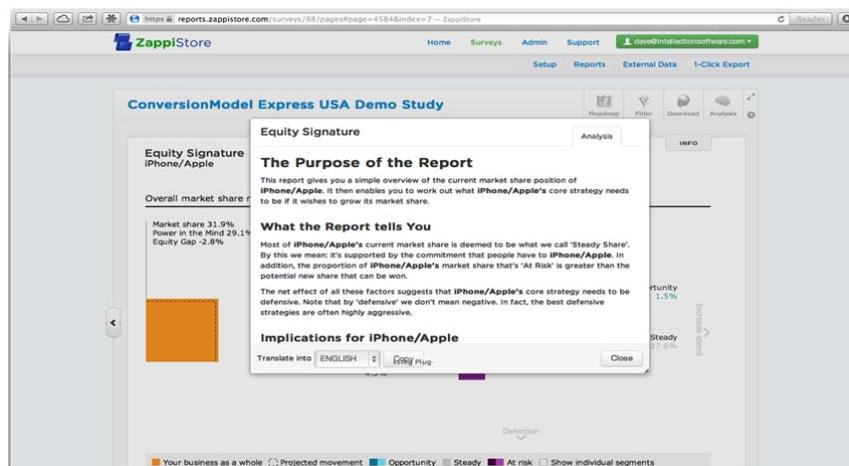
The tool goes on to cover all of the standard market research metrics:

- Ad likes and dislikes
- Emotions created by the ad
- Messaging attribute testing
- Recall and branding, credibility, purchase intention
- Viral likelihood

All of these results are shown in this online reporting tool that can easily be manipulated to filter results on key target groups or to create bespoke presentations as required. These presentations can be immediately downloaded into PowerPoint or as a PDF. So client side researchers can go into a meeting with their marketing clients within minutes of receiving the results.

Figure 2 is another screen shot showing some of the power of the machine learning. This comes from a test survey (non-confidential) conducted using one of the branded research company tools on the site, the TNS Conversion Model™, and shows the extent of the strategic recommendations generated through the reporting engine. Again, this analysis is entirely automated and happens within seconds of the data hitting the system.

FIGURE 2.



In order to create this level of analysis the system has to digest and model the thinking of the person who designed the product, here either the online research provider's ad test or the TNS Conversion Model™. The process of distilling these potential interpretations and human analysis of the data takes time but once it is done the system can replicate the

analysis, for multiple market scenarios, instantaneously. This allows clients to access the best human thinking but at machine speed.

Because the entire system is automated, it is possible to complete a project from start to complete report / presentation with analysis very quickly - in some markets in a matter of hours. Some markets and audiences take longer due to the lack of availability of sample but the results come in significantly quicker than using traditional tools.

In theory this means that an advertising creative or product developer could have an idea in the morning, test the idea over lunch and then come back to the idea informed with real consumer insight and strategic recommendations, in the afternoon. This is the promise of agile research. Yet, it was unclear if the system and sample would work in Southeast Asia, and it was also not clear that the team at Coca-Cola could overcome natural internal barriers to adoption.

### THE VALIDATION APPROACH

Rather than demanding solutions be tried and tested already, Coca-Cola sought to invest in the test-and-learn process along with their partners. While there was much excitement about the potential for the new tools to inform decision making and an iterative approach to development, especially in the early stages, a tempered test-and-learn approach was essential. Creating confidence in the new framework meant beginning and validating in a small number of countries.

Specifically, Coca-Cola established a plan to develop new benchmarks with these agile tools where none previously existed, and leveraged the expertise of the online sample agency to assure the samples were working optimally.

Testing centered initially on advertising copy testing. Concurrently, parallel testing with traditional methods was used to establish the validity of experimental approaches. At the same time The Coca-Cola team created new processes that integrated the new approach and findings while securing the support of key marketing and IMC leaders in early usage and learning.

#### Generating hypothesis, testing validity

To kick start the process the team came up with a series of hypothesis of factors we might encounter and which may cause this approach to *not* work that we could then directly test. By starting in this way we knew we would learn more quickly and have confidence in the results where our hypothesis were shown to be FALSE and also that we could answer any questions or concerns that the business may have in order to ensure that the approach, if it worked, would take off.

So our initial hypothesis grouped into three key areas:

1. Would the business make the same decision from a traditional approach compared to the new agile approach?
2. To what degree are results directly comparable or directionally comparable between these traditional and new approaches?
3. Are there specific countries or audiences that definitely do not work or where the results should be treated with extra caution due to the movement from a face-to-face to online methodology?

In order to answer these questions we designed a series of research experiments using existing Coca-Cola research conducted in traditional ways and directly compared them to new test results using online sample.

#### Validation 1 – Would the business make the same decision?

The Coca-Cola Company already uses ad testing from a leading research vendor who has very strong credentials and validated approaches in Asia. We knew that their tests were a great benchmark to determine how the new approach would compare.

The first level of validation included doing a series of tests on TV advertising executions that had already been tested using the existing approach. These tests were conducted across Thailand, Indonesia, Vietnam and The Philippines. The two methodologies were:

- Face to face, central location testing in two locations per country
- Online sample across the whole country using the agile tools

We must then of course concern ourselves with mode effects outside of the sampling itself. The online mode is self-completion which has the distinct advantage of reduced social desirability bias. Social desirability bias comes into play when the respondent wishes to please the researcher or conceal from the researcher true feelings that are not the social norm. It is surprisingly common and very much related to culture in terms of its direction and strength. As Dillman (2009)

notes “[it] even operates at a threshold far below what one thinks of as a sensitive behavior”. Other mode effects include the way in which scales are responded to (more use of the centre points), or brands are recalled (more brands recalled without interviewer pressure to ‘move on’ with the survey) or how often ‘don’t know’ is used (respondent happier to admit ignorance if online).

The questionnaires were also different and so a direct comparison of specific questions was not possible. However, the areas covered were similar enough to allow some basic comparisons about which ads perform well or poorly.

However, the purpose of the proposed tools at this stage was to be an early quantitative review that would be superior to focus groups in revealing the potential strength of an ad in its early stages. Exact replication of the face-to-face results was not required as long as it correctly guided the business. The Coca-Cola team reviewed the results to see if the business would have made the same decisions about what to do next that ultimately came from the industry standard tool. To add a more quantitative validation we looked at the rankings of the different ads on a similar metric for which we picked ‘enjoyment’. This was a liking metric within the Agile Tool creative test and an ‘engagement / enjoyment’ metric within the face to face survey but it seemed to reflect the same basic ad response mechanic. The results shown in table 1 are disguised in terms of the ads shown:

**TABLE 1. INDONESIA**

	Indonesia	
	Agile Tool Ranking	Face-to-Face Ranking
Ad 1	5	5
Ad 2	4	4
Ad 3	1	1
Ad 4	3	3
Ad 5	2	2

Indonesia had a perfect correlation between ad rankings suggesting that using the Agile Tool results for early iterative learning would indeed be practical and would be in line with later tools in the TCCC toolbox. Our small test suggested that the online sample would make little difference to the validity of results and made us feel hopeful for other countries. However, Indonesia has a larger panel base within SSI which made it more likely that the sample covered would cover the different demographic sub groups.

Thailand was the next market reviewed (see table 2).

**TABLE 2. THAILAND**

	Thailand	
	Agile Tool Ranking	Face-to-Face Ranking
Ad 1	4	3
Ad 2	2	4
Ad 3	1	1
Ad 4	3	2

Here the results are less conclusive. The best ad came out across both methods and there were similarities in direction but the absolute fit was less strong. We now needed to understand if this was a sampling problem or a cultural one. Clearly we have more to learn in Thailand and must be careful until we figure out more, despite the exciting results from the Indonesia examples.

Results for Vietnam posed more problems as we did not have access to the enjoyment scores from the previous testing methodology. Instead we had a composite rating and we had seen in the other data that the composite measure correlated less well than enjoyment which was more closely aligned in terms of questionnaire (see table 3).

**TABLE 3. VIETNAM**

	Vietnam	
	Agile Tool Ranking	Face-to-Face Ranking
Ad 1	5	6

Ad 2	1	1
Ad 3	2	3
Ad 4	4	4
Ad 5	6	2
Ad 6	3	5

Having said that we were working with a composite measure, we still saw strong correlations between the rankings with the exception of Ad 5 which scored very highly on the composite measure but poorly on the Agile Tool enjoyment metric. Looking at the findings from the Agile Tool survey showed that the ad in question was, in fact, strongest on some other measures such as 'Relevant' and 'Behaviour change' and in the top 1 or 2 tests for 'Purchase intention', we just saw a high degree of indifference to the ad on our liking measure (see table 4).

**TABLE 4. THE PHILIPPINES**

	The Philippines	
	Agile Tool Ranking	Face-to-Face Ranking
Ad 1	5	5
Ad 2	2	2
Ad 3	3	1
Ad 4	6	6
Ad 5	4	3
Ad 6	1	4

Similar to Vietnam we found broad correlations between results in The Philippines with the lowest rated ads scoring poorly on both tests and similar results for most of the others except for Ad 6 where again we saw strong results for the face to face test and in the Agile Tool study the results were weak in the overall liking metric. Again we looked in a little more detail and saw the same issue with the ad scoring much better on the emotional metric within the face to face test than it did on some of the branding and persuasion questions.

This seemed to back up the qualitative view that the recommendations from the reports would send you in a similar business direction even if the specifics of the cross ad test were more difficult to make.

So at this stage we had promising results and a belief that an agile research approach using the Agile Tool technology combined with the research agency panel in the early stage testing approach would be practical. Obviously there were still concerns, particularly with respect to panel representativeness, but the initial results were certainly encouraging.

#### **Validation 2: Copy Testing: Within population responses**

The second stage of this test was to look in more detail within the online results to see if there were some inherent biases within the overall target audience that could derail the usage of online samples.

It is of course entirely possible for online research to be conducted in any country where an online access panel exists. By definition the population the online access panel can be projected to is the online population, and this is rarely the target population of choice. We must acknowledge then that there is a coverage error in using an online access panel, however the quotas are defined, and that this error will be greater in countries with lower internet penetrations. That said, face-to-face research is often done within limited geographies to contain costs, so both sources may not represent the overall population.

The current internet penetration and online panel sample structures within each market is shown in table 5:

TABLE 5. THE PHILIPPINES

Country	Population	Internet Users	Internet Penetration	Online Research - Representative	SSI Panel Size
Malaysia	30,073,353	20,140,125	68%	Possible	68,043
China	1,355,692,576	642,261,240	50%	Tough	981,348
Vietnam	93,421,835	41,012,186	50%	Tough	60,201
Hong Kong	7,112,688	5,751,357	76%	Yes	106,834
Singapore	5,567,301	4,453,859	87%	Yes	106,012
Taiwan	23,359,928	18,687,942	80%	Yes	134,210
Philippines	107,668,231	44,200,540	49%	Tough	154,153
India	1,236,344,631	243,000,000	19%	No	237,561
Indonesia	253,609,643	71,190,000	21%	No	402,778
Thailand	67,741,401	20,100,000	51%	No	134,964
Myanmar	55,746,253	668,955	1%	No	0
Cambodia	15,458,332	927,500	6.00%	No	0
Laos	6,803,699	850,425	12.50%	No	0

Source: EIU (internetworldstats for Cambodia and Laos) & SSI

The internet penetration figures (from The Economist) also show very rapid growth in online penetration in some markets. For instance Thailand has gone from 24% in 2011 to 51% in 2014, similarly Vietnam has gone from 35% to just over 50%, The Philippines from 31% to 49% and Indonesia, the laggard of our markets in terms of penetration, has still managed to almost double in that time period from 12% to 21%.

Given this rapid rise in online penetration the use of online research as the industry standard (as it is in the West) is getting closer and closer. Online panels sizes and sample distribution will certainly increase rapidly given this change in online penetration and we would expect more and more companies to be benefitting from the cost, timing and quality advantages that online offers.

The panel sizes shown clearly enable fast turnaround projects in markets such as China, Taiwan and potentially India and Indonesia. However, panel size and panel strength are not necessarily the same thing and even within panel size the representativeness of the panel (or lack of it) can undermine a panel.

What we hoped was that the panel we used in each market; while not necessarily ideal in terms of market internet penetration, panel size or panel representativeness; would still enable good enough research to enable businesses to gain from the benefits that agile research promises.

As is clear from the table we were looking at markets, particularly Indonesia and Thailand, with moderate internet penetration, but huge internet populations of over 20 million people. Although the online population will not be genuinely representative of the market overall, it represents a huge base of people. Therefore, when looking at the stats in more detail we wanted to see what types of biases we were potentially opening ourselves up to. One hypothesis was that young people would be well represented but that older people were less likely to be online. This is born out in the stats from Indonesia shown in table 6.

TABLE 6.

Age	Penetration (%)	Online Population Distribution (%)
15-24	51	43
25-34	33	27
35-44	26	19
45-54	15	8
55+	6	4

Source: Comscore

If the target population is under 35 years then online gives very good coverage in most countries – better in many instances than landline telephone or face-to-face. Other distinct biases are towards higher incomes and the better educated (and often these are positively correlated as well). Again reference to the target population is important. In developing economies often the poor and ill-educated are, essentially, outside the consumer economy entirely and living at a subsistence level. The reality then can be that online is actually more representative of a “general population” when that general population is in fact the consumer economy.

**TABLE 7. PERCENTAGE OF EMPLOYED POPULATION AGED 15 YEARS AND OVER USED INTERNET BY OCCUPATION**

Professionals	88.50
Clerks	75.00
Technicians and Associate Professionals	74.60
Legislators, senior officials and managers	53.80
Service workers and shop	21.00
Craft and related trades workers	13.60
Plant and machine operators and assemblers	11.20
Elementary occupations of sales and service	4.20
Skilled agricultural and fishery	3.80

*Ministry of Information and Communication Technology Thailand 2012*

As we see from table 7, if your target audience is the more educated and professional audience then a standard online sample looks to be potentially very representative. And, given the rapid penetration of Internet access in Thailand, current concerns about representativeness of the more manual functions will continue to ease, although genuine caution is currently needed.

So, this does suggest that conducting online work among those aged under 30 years who are well educated can be representative. But what about other target audiences? Does it mean that brands who want to include over 30 year olds should not go online?

The traditional Coca-Cola target audience for face to face studies had been:

- Past week carbonated soda drinkers
- Ages 12-49 years
- All social economic classes
- Males and Females equal split

When we started running studies through the new system the target group was expanded to account for the relatively smaller online panels and became:

- Past month carbonated soda drinkers
- All ages 18+ years
- All social economic classes
- Males and Females split 60:40 given the larger numbers of males on the panels

Again this makes the studies not directly comparable and when looking into the panels in more detail it became clear that there were specific issues:

- Men were significantly more represented than women, sometimes up to 80:20
- Very few lower social groupings were on the panels due to lower internet penetration

**TABLE 8. POPULATION 5 YEARS OF AGE AND OVER USING INTERNET IN THE LAST 3 MONTHS BY HIGHEST EDUCATION LEVEL ATTAINED, 2013 (%)**

None	5.36
Primary School	14.14
Junior High School	21.12
Senior High School	37.15

*Source: Dept of Statistics Indonesia 2013*

However, given the recent interest in James Surowiecki's *The Wisdom of Crowds* we decided to test how large an influence this under representativeness of specific target groupings had. On the new tests we did have both men and women, we had broad age ranges and some lower social groups and so we decided to see how their reactions to the advertising stimulus varied. If indeed we found very large differences between groups then we would need to be very nervous about using online samples that have in build representativeness issues.

Below is some of the data from our studies to illustrate the differences we found across audiences. The data is from Thailand and shows the mean scores for 'liking' of a couple of ads chosen at random across different audience segments. We selected a range of issues to investigate to determine how much potential bias was created (see table 9).

TABLE 9.

	Income		
	Low	Middle	High
top 2 box	26%	27%	31%
mean	7.41	7.43	7.64

*Clearly some bias towards Higher income audiences liking this ad more*

	Panel Tenure				
	0-3 months	4-6 months	7-9 months	10-12 months	more than 1 year
top 2 box	29%	29%	37%	30%	29%
mean	7.41	7.5	7.97	7.7	7.68

*Perhaps some bias towards those who have been on panels longer*

	Age				
	<25	25-34	35-44	45-54	55+
top 2 box	28%	28%	28%	30%	25%
mean	7.39	7.48	7.47	7.72	7.49

*No obvious bias regarding ages*

	Gender	
	Male	Female
top 2 box	28%	29%
mean	7.4	7.56

*Some bias identified between males and females*

	Education					
	Low	2	3	4	5	High
top 2 box	18%	19%	16%	25%	23%	23%
mean	7.24	6.97	7.06	7.14	7.16	6.83

*No obvious correlation between education and reaction to the ads*

So we did find some biases but even the most extreme bias only moved top box results by up to 4% from the mean and most of the target biases lead to results being only 2%-3% different from the average. These differences by target group seem to be small enough not to distort the overall findings for a study. This does suggest that there is something to the Wisdom of Crowds idea and that as long as we are broadly representative then worries about under-representing specific sub-target populations is of less concern than we researchers typically think.

At this stage we think a pause may be required. We have seen that:

- Directionally, comparing face to face methods to online samples seems to generate ‘similar enough’ results that businesses would make the same business decision off the back of them – at least for advertising evaluation studies
- Online panels in ASEAN do have inherent biases BUT these biases may not impact research results as much as we typically fear

So the preliminary results seem to be favourable for the use of the new approaches in early stage ad testing, which enables faster, cheaper research to be conducted. This in turn enables businesses to take a more iterative approach to product or communications development, testing things more often and helping make consumers much more central to the decisions businesses make.

### Validation 3 – Direct, like for like research question validation through tracking

Coca-Cola has a continuous equity tracking study going on across a number of markets. Could an agile tool work there for speed and efficiency? We decided to see if we could use this tracker as a direct comparator and so quickly put an online version of the tracker into field using the same sampling criteria and question structure as the current tracker – therefore directly comparing an online approach with a face-to-face tracker conducted in two locations in each market.

Again we knew that there would be differences caused by:

- Comparing face to face methods with online
- Comparing data that covers the entire country (online) with data from just two central locations (offline)
- A move from advertising testing (asking about liking something) to brand tracking (often asking about actual behaviour or recall where you might expect greater interviewer bias)

## SWITCHING METHODOLOGIES

When you switch from face-to-face or telephone to online one thing is certain; your data is going to change. For many, particularly those involved in tracking research, this can be traumatic. Much faith has been invested in the veracity of the existing tracker which makes it hard to believe some new truth arriving from some new methodology. And yet a deeper understanding and appreciation of the dynamics involved in survey mechanics can help make sense of how two different truths can co-exist.

Switching modes of interviewing essentially involves three changes:

1. Sampling changes
  - from the total population to the online population. This represents a non-coverage error
  - from probability methods to access panel methods
2. The mode changes from an administered mode to one of self-completion
3. The questions change from a heard “conversation” to a read set

To understand the impact of each requires a different set of comparative measures that may go above and beyond simply comparing two data sets.

Our data sets consist of three distinct experiments:

- Experiment 1 – A straight transposition of a tracking survey to online. Comparing existing offline to online
- Experiment 2 – An adjusted offline survey that includes questions relating to respondents online access. Comparing new offline to online
- Experiment 3 – An innovative approach in which respondents, recruited face-to-face, are either administered the questionnaire as normal or are given access to a tablet and invited to fill in the survey by themselves, in private. Comparing administered surveys to self-completion surveys

Experiments 2 and 3 are still in field at time of going to press but will be presented at the conference itself.

## SAMPLING

To understand the impact of the non-coverage one needs to hold constant the mode of interviewing and the question sets and examine only the differences attributable to having access to the internet. The potential for impact is, of course, essentially mathematics. If the overall internet penetration is 70% then it is hard to envisage a case where the absence of 30% of the population can make a material difference to the overall data, given that the online and the offline population rarely hold opinions that are diametrically opposed to each other. Large biases can occur when the subject matter is strongly related to being online. Measures of usage of bricks and mortar banks, for example, might be underestimated using only an online sample.

Bias should occur then only if the subject matter of the survey is correlated with being online (or not) and material effects only occur if the non-coverage is large.

Non-coverage is high in Indonesia (79% without internet access), because of a significant rural population. Yet, among the urban populations coverage will be moderate. Non-coverage is moderate (about 50%) in Thailand, Philippines and Vietnam. It is relatively low in Malaysia (32%) and Singapore (15%).

Sample selection in an online setting is much more positively self-selecting (particularly when using survey invitations) (Cape, 2014) than in traditional methods. There it is the persuasiveness of the interviewer that can determine who takes part. In face-to-face quota methods the active choice of the interviewer of who to approach can also play a part.

Since almost all online surveying is done using access panels the researcher also needs to concern him/herself with what happened in panels. It has long been established that tenured panellists respond differently to newer panellists (Gailey et al 2008), and that the different original sources of panellists can lead to different survey outcomes (Willems et al, 2006).

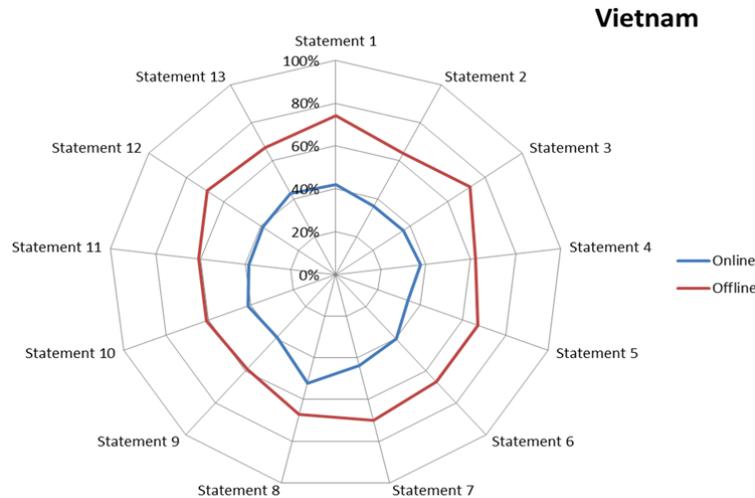
Control of these latter variables over time is paramount since there is no essential truth that can be held fast to. Most sample providers keep constant their sources, measure tenure and utilise standard survey invitations or, preferably, survey routers.

Non-coverage however can be measured. This is simply done by holding constant the sampling methods, the mode of interviewing and the questions themselves and then analysing the two groups – online population vs. total population.

Usually this is in conjunction with an analysis by age since the online population is heavily skewed by age in most unsaturated markets.

Experiment 1 exhibits what appears to be a sample effect on brand image.

FIGURE 3.



The numbers are different but the pattern relatively consistent. Perhaps the online population, younger and more adventurous, better educated and richer, has a different view to the total population? Assuming both data sets to be true and given an internet penetration of 50% among 18-49 year olds we can assess what the brand image of the offline population must be in order for them both to be true (see table 10).

TABLE 10.

	Online population	Calculation of Offline population	Total population
Statement 1	42%	106%	74%
Statement 2	36%	92%	64%
Statement 3	36%	108%	72%
Statement 4	38%	86%	62%
Statement 5	34%	100%	67%
Statement 6	40%	94%	67%
Statement 7	44%	96%	70%
Statement 8	52%	82%	67%
Statement 9	39%	80%	59%
Statement 10	41%	80%	61%
Statement 11	39%	84%	61%
Statement 12	39%	100%	69%
Statement 13	43%	92%	67%

Is it credible that the offline population thinks so differently? It seems unlikely but is the only explanation if we assume no mode effects or questionnaire effects. Certainly the offline data does not vary by age within Vietnam, and recall that Internet usage is almost twice as high among 20-29 year olds than for 40 to 49 year olds, we should see some effect here if the above table were true.

Only a true like-for like comparison, as we have in our other experiments can show the true sampling effect.

### MODE EFFECTS

Since the acceptance of online research as a 'fit for purpose' methodology there has been a renewed interest in mode effects, effects that had been long known to academic researchers who had continued to utilise self-completion postal questionnaires long after the commercial world switched to telephone and face-to-face. The effects come from the presence of the interviewer. Respondents, as humans, want to portray themselves in a good light and will give socially desirable answers. These effects are strong and pernicious and can occur in the strangest of settings and can switch direction as one changes culture (Wichers, 2006). The direction and particularly the strength of the bias is therefore

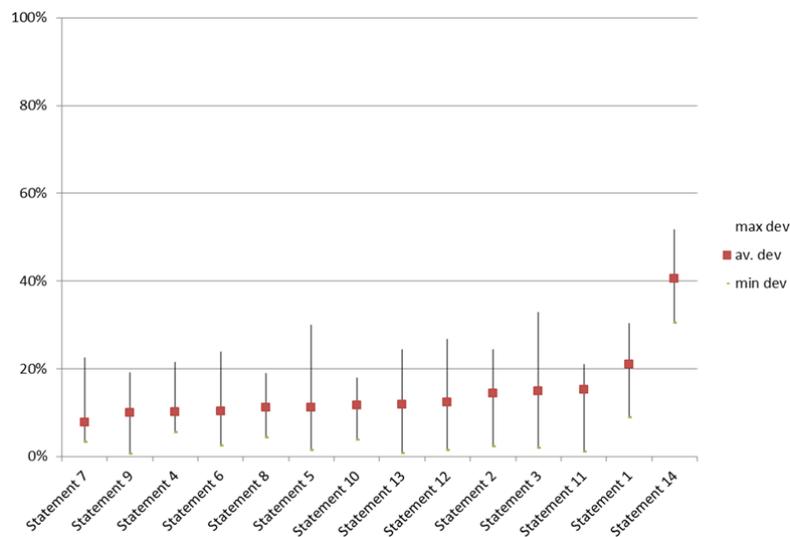
difficult to predict, experiments into calibrating social desirability bias are only now starting to emerge (Gittelman, 2015). Similarly the ‘desire to please’ is strong. The interviewer can be seen as representing the sponsor, and the respondent wishes to please the sponsor by being positive about the brand, by displaying high levels of customer satisfaction or perhaps by claiming consumption of the brand at all.

In order to measure mode effects one needs to hold the questionnaire and sampling constant and vary the mode.

The presence of socially desirable answering is clearly seen in the brand imagery section within the Experiment 1.

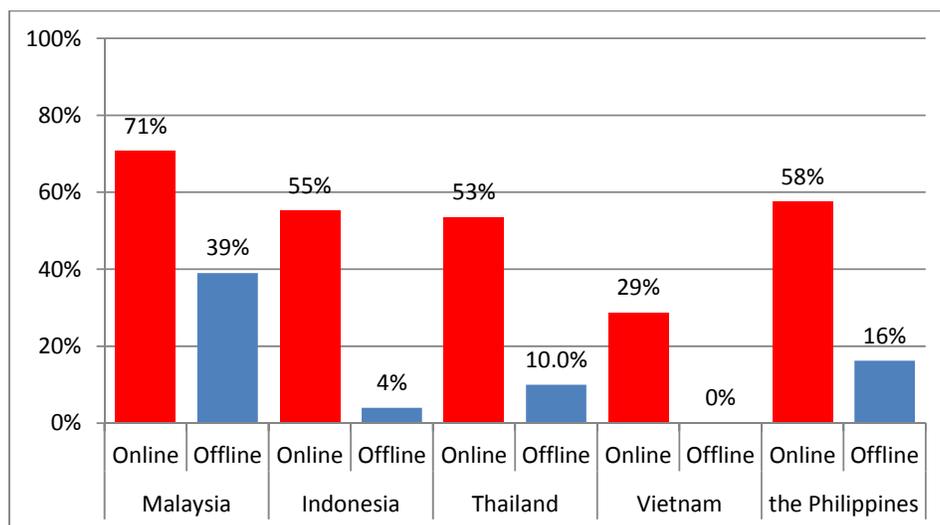
Figure 4 shows the deviations between the offline and online data collected for one brand across all the countries. The average deviation is shown along with the minimum and maximum.

FIGURE 4. BRAND IMAGE: OFFLINE VS. ONLINE



One item in the attitude battery was very different and very clearly subject to bias – Statement 14. This statement asked about a social perception of the brand. The need to treat the brand kindly in front of the interviewer is clear and the deviations of online compared to offline are huge. Most respondents consumed the brand and were hesitant to mention a negative social perception. Online respondents did not show the same social bias. (See figure 5.)

FIGURE 5. SOCIAL PERCEPTION QUESTION



That this is not a sample effect can easily be proven by mathematics. Malaysia has 68% internet penetration. If all the offline people in Malaysia do not believe in the social perception then 71% of 68% of the offline sample must believe it. This results in a theoretical minimum of 47% believing it unhealthy in the offline sample – where 39% is recorded. Notice also that the strength of the bias is different per country. In Indonesia for example the theoretical minimum is three times lower than the actual recorded in the offline survey!

The presence of the interviewer has other, un-unintended consequences also. The interviewer controls the pace of the interview and this has an effect on, for example, the number of answers given in spontaneous questions – higher in self-completion where the respondent has more time to search for that elusive last brand. The interviewer engages in a dialogue with the respondent in the open questions, exploring and clarifying whereas in online responses are thin and represents first response only (Pete, 2015).

### QUESTIONNAIRE EFFECTS

Again the interviewer's role may be crucial here. In both telephone and face-to-face interview situations the interviewer is, in theory at least, a simple conduit for the questionnaire as written. However, it would take a heart of stone not to help the respondent as they struggle with the meaning of questions, and few do not. This reliance on the interviewer to clarify ambiguity and give further sense to the questions can lead researchers to be somewhat lazy in their approach to questionnaire design. Simply transposing a poor questionnaire design to an online setting leaves the respondent on their own and they must answer the question as they see fit. Appreciation of the effects of poor questionnaire design and poor question wording has been a long uphill battle in online research (Pete et al, 2008).

Other questionnaire effects are less surmountable. Online respondents use scales differently to interviews administered offline. More use is made of the centre points in online compared to offline. This is entirely due to the mode of administration and the online questionnaire cannot be made to mimic the offline. An offline questionnaire can be made to mimic an online survey (Wine et al 2006) but this is of little comfort to the researcher transitioning their offline tracker to online.

Online respondents exposed to a grid style of question, used typically in online studies to save interview time, will also answer differently as the items are now in a comparative framework (Gittelman, 2015). In addition there is no control over which direction the grid is completed (Pete et al 2008) or the comparisons made. And, of course, there is the ubiquitous social desirability bias.

In all our experiential questionnaires no scaled questions were asked.

Questionnaire effects can be seen by experimenting with different question wordings within an online setting.

Whilst we have no experimental treatments expressly looking at question wording we do find an interesting anomaly in Experiment 1. The survey uses 'standard' category names that may or may not be well understood by the respondent. One category in the offline study is universally consumed. Anecdotally this is believed to be true and raised no comment. In the online survey the headline numbers are considerably lower. That this is a questioning effect is evidenced by the finding that more people drank brands from the category than drank the category itself. Adding in the relevant brand consumptions into category usage increased category usage by 25%. Only in the case of Thailand does the raw online data even begin to approximate the offline. A back translation of the Thai questions reveals that the local Thai survey does not use the standard category name but instead translates it to the vernacular name! (See table 11.)

TABLE 11.

	Malaysia	Indonesia	Thailand	Vietnam	the Philippines
Online category usage	64%	64%	83%	69%	69%
Online category usage + Brand	80%	81%	92%	88%	86%
Offline category usage	100%	97%	97%	100%	100%
Uplift of + Brand	+25%	+27%	+11%	+28%	+25%

The question may remain: why do not all Thai's say they use the category when they actually do? The answer goes back to survey dynamics. The respondent knows they are in an experiment, not in a conversation with a friend. Being asked a question, the answer to which is universally Yes, may appear to be a trick, something that isn't. Just as at least 10% of respondents will claim to use brands that do not exist, it is extremely difficult to get respondents, particularly online access panel respondents who see a lot of questionnaires, to admit to breathing, eating, or even using the internet whilst they are filling in a survey online.

It is hard to conclude other than that some degree of help is being given to respondents in an offline setting to enable them to reveal their truth. Therefore the online questionnaire must give the same help and guidance, extant – not just be a straight adaptation of the offline questionnaire as written.

### DIRECT COMPARISON SUMMARY

Faced with the various forces pushing and pulling the data, the importance to the client of consistency and the assumed validity of the existing data, it is tempting for the researchers to either resist the new methodology or mistrust the online respondents themselves.

Online research, however, has repeatedly been demonstrated to be consistent and fit-for-purpose even if not technically projectable.

A starting point of believing that the respondent is consistently giving their truth is necessary in order to disentangle the various effects. Once understood, researcher and client are in a better position to accept a discontinuity in their data and ‘move on’, taking advantage of everything online research has to offer.

### VALIDATION SUMMARY

Clearly making a move into online research in markets with low internet penetration is going to be a risky venture. We wanted to try and look at the implications of this in detail and see how risky that move would be and what impacts we could potentially predict and therefore mitigate against when making the move.

So where have we gotten to so far, we believe that we can say:

- It appears that results of a perceptual nature, in this case advertising reactions / liking, have been shown to be consistent enough between an online and offline methodology to validate the use of online and therefore the new agile approach for ad testing.
  - We would hypothesise that this would also be true for other types of testing such as product or pack concept testing but again this needs to be tested.
- However, when looking at more behavioural metrics and absolute scores there is still significant differences seen between online and offline. So even if they are directionally similar, we require more study into the effects of this switch before we recommend moving to an online tracking approach.

### IMPACT ON COCA-COLA - WHAT IMPACT DID THIS HAVE ON BUSINESS THINKING?

The results of this work were highly encouraging and deeply instructive. Given the fact that business moves faster than ever before, marketers need to make thoughtful decisions more rapidly, but with confidence. And given the iterative nature of the work to be done market researchers need to add additional tools and flexibility to help them make those decisions in a better way.

As was hoped, the work gave the Coca-Cola teams in Research and Marketing confidence in using the automated tools and online sample for early stage testing. The anomalies we observed highlighted some issues we must explore further to improve our understanding of the tools and approaches used.

When looking at the timings we have seen for these new tools we can see how quickly studies can be delivered:

TABLE 12.

	Shortest	Average
Japan	56 minutes	19 hours 17 minutes
China	1 hour 58 minutes	2 days 1 hour
Singapore	3 hours 21 minutes	11 hours 52 minutes
Thailand	23 hours 41 minutes	4 days 19 hours
Malaysia	1 day 12 hours	3 days 21 hours
Vietnam	1 day and 15 hours	3 days and 19 hours
Indonesia	3 days	4 days 2 hours

Here we see the shortest time to complete a project in each of the Asian markets for The Coca-Cola Company and also the average time taken. We expect these timings to shorten as panel sizes increase within the region. The timing depends on two factors:

1. Panel efficacy (size and engagement) in the market in question
2. Target audience definition used on the project

If the target audience is Nationally Representative then projects will fill much quicker than if we are looking at 18-34 year old males who live in Jakarta and like Britney Spears! We also learned that more generalised audience definitions do not necessarily significantly impact results. A general audience may well be preferred in many situations.

Fast results like these created a new opportunity for The Coca-Cola Company to improve the creative development process and to deal with unexpected results. Creative timelines are always very aggressive. Testing and refining top-quality advertising with the required timelines in one country can be a challenge. Doing it well across multiple countries multiplies the difficulty. The quick turnaround that has been experienced so far allows Coca-Cola to get a rapid quantitative read on the quality in the early animatic stage, giving back a very critical four to seven days to allow creative teams to refine and improve the work. Furthermore, when a problem with creative emerges in a particular country, the ability to refine, test quickly, refine, and test again has allowed The Coca-Cola teams to get back on track more quickly. Coca-Cola is now beginning to realize the promise of Agile research that was hoped for when this initiative began.

Further, this initiative changed the thinking of research associates in the Coca-Cola team. The online research world in Southeast Asia is understood in ways that were not comprehended in the past. The associates were always willing to explore, within boundaries, but applying Test-and-Learn to our own research tools is now a stronger part of the thinking. Those boundaries have been stretched in a positive way, and it has spread in a positive way to the Marketing team as well. More and more Coca-Cola Company research will be done online in Southeast Asia using fast and efficient agile research tools.

## CONCLUSION

Creating a win-win-win is not too easy to do. It requires the courage to explore, invest time and resources, and to collaborate for mutual benefit. Yet, there was so much that was learned from this effort.

1. Researchers in Southeast Asia can use online research with reasonable confidence in some types of studies, as was shown in the early testing examples. Although online research is far from perfect in Southeast Asia, new tools like these can have a meaningful place today.
2. More work is needed to understand some anomalies in the findings. However, the causes of many apparent gaps became obvious because of understanding from prior global studies. There is great wisdom in connecting and collaborating with those familiar with global work. Such efforts can help you to not reinvent the wheel (example: when comparing self-complete and assisted-complete approaches).
3. There is much more to learn as researchers in Southeast Asia build on traditional ways of doing research. Forums such as ESOMAR create an ideal climate to explore, learn and share.
4. Because ZappiStore and SSI were willing to join The Coca-Cola Company research team to collaborate on this learning journey, the reality of using fast and efficient tools has been significantly advanced. What may have taken several years to drive has become a current reality.

For us this means there is more work to be done, more experiments to be conducted and more validations to work on! However the promise of Agile Research is now real, we can conduct some very important types of studies from beginning to end in hours if not days, no longer weeks or months. This is already changing the way The Coca-Cola Company utilises consumer insight to inspire great creative work.

The Coca-Cola Company believes that researchers need to make continued movement towards a more agile world as we develop our suite of tools. The benefits of doing so are very clear and the dangers of not moving in this direction are that the insights world may become increasingly divorced from the fast paced world of marketing and the businesses where we work. Collaboration with like-minded people will be critical to move such thinking forward.

## BIBLIOGRAPHY

- Cape, Pete, Lorch Jacqueline, Piekarski Linda, Back to Skool?, 2008, Melbourne, AMSRS
- Cape, Pete. Quota Controls; Science or merely Sciencey? 2011. CASRO Online Conference, New Orleans.
- Cape, Pete. 2013. Do I have your full attention. San Francisco, CASRO
- Cape, Pete. 2014. Interest Bias. Koln, GOR
- Cape, Pete. 2015. Using video to re-invigorate the open question., Nashville, CASRO
- Dillman, D.A. (2007) Mail and Internet Surveys: The Tailored Design Method, 2nd Ed., New Jersey: Wiley
- Dillman, Smyth & Christian; Internet, Mail and Mixed Mode Surveys, The Tailored Design Method; Wiley 2009, p 313
- Gailey, Ron, Teal David, Haechrel Eric. 2008. Sample Factors that Influence Data Quality, WaMu
- Gittelman Steve. 2015. Accounting for Social Desirability Bias: A Model for Calibration of Direction and Magnitude, 2015, Nashville, CASRO

Surowiecki, James. The Wisdom of Crowds

Willems Pieter, Van Ossenbruggen Robert, Vonk Ted, 2006, "The effects of panel recruitment and management on research results", Dutch online panel study NOPVO, Barcelona, ESOMAR

Wichers Bart, Zengerink Evelien, 2006, "It's the culture, stupid! A cross-cultural comparison of data collection methods", Amsterdam, ESOMAR

Wine, Cominole, Heuer & Riccobono, Challenges of Designing and Implementing Multimode Instruments, 2006, RTI International

## **THE AUTHORS**

Ron Gailey is Director of Insights, The Coca-Cola Company, Thailand.

Stephen Phillips is CEO, ZappiStore, United Kingdom.

Pete Cape is Global Knowledge Director, SSI, United Kingdom.