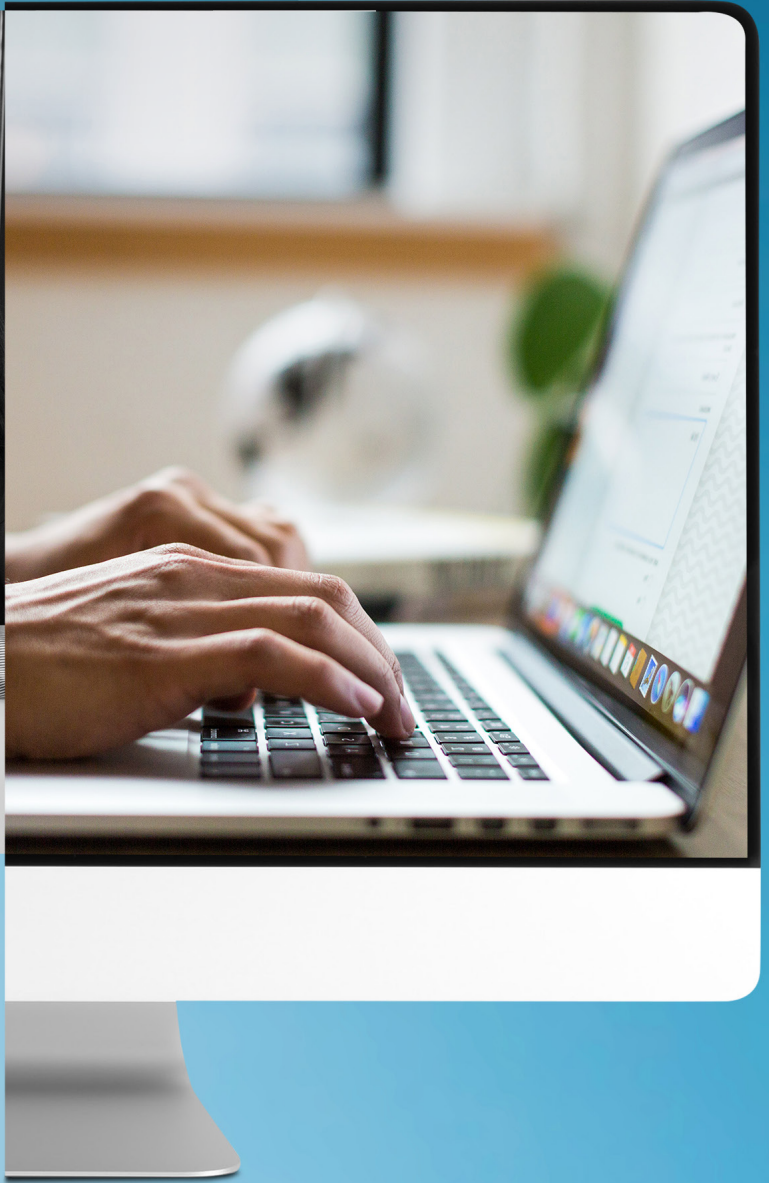# A NEW WORLD

What to consider when transitioning a
research program from offline to online

By **Sara Wilkinson and Leah McTiernan** | March 2020

IPSOS
VIEWS

GAME CHANGERS   Ipsos

INTRODUCTION

ADVANCES IN TECHNOLOGY AND INCREASED ACCESS TO THE INTERNET HAVE MADE ONLINE RESEARCH MORE VIABLE GLOBALLY.

With the pricing and cycle time benefits that online research provides, transitioning programs from offline to online should be a key consideration. However, the decision to change is not always an easy one, as change often brings disruption to data trends.

Trends disruption can be a barrier to transitioning existing studies from offline to online. Because Ipsos truly understands the data differences, we can help clients prepare for these potential differences. Over the last 20 years, Ipsos has amassed a wealth of experience in transitioning research programs, first with extensive parallel testing programs in North America and Europe, and now spanning all regions.

The following outlines key learnings from our experience to help clients make informed decisions, including how and when to transition and what to expect.

WHY DOES TRANSITIONING FROM OFFLINE TO ONLINE BRING DIFFERENCES?

Two broad factors are at play: the population that the survey method reaches and the presentation that the survey method offers.

POPULATION DIFFERENCES

Most studies follow a convenience sampling approach led by the data collection method. As such offline and online select people for surveys in different ways. None is inherently right or wrong, but each is likely to produce different results from one another.

Offline is a broad data collection category and includes face-to-face and telephone interviewing. Some face-to-face studies recruit people based on postal addresses, offering

extensive geographical coverage of the full population. Others are clustered around interview points or central locations that limit or prohibit the coverage of certain geographies, often rural areas. Telephone studies achieve a geographically dispersed coverage but are limited to only on consumers with a landline or mobile phone, for which access has become increasingly restricted.

When it comes to online, an online survey can of course reach only those consumers that are online. In countries where large parts of the population are not online, an online approach limits or excludes consumers that are less likely to have access to the internet, such as older people, people with a lower socio-economic status and people living in more rural areas. Within online, differences in the types of people selected for a survey can stem from where on the internet people are recruited, particularly if certain types of internet sites, such as social media, reward or gaming sites, direct people towards a specific survey or survey panel.

SURVEY PRESENTATION DIFFERENCES

How somebody answers a question is influenced by how the question is asked.

The same question can be interpreted by the same person in different ways depending on the setting. In offline, the meaning conveyed by an interviewer's tone, enunciation and reading pace influences the respondent. In online, how the question looks on screen, survey layout and images have an impact.

A transition from offline to online often implies changes to questionnaires themselves, to make them suitable for the online method. Questionnaires are often shortened and simplified to fit with online survey-taking devices, from larger screen sizes for PC and laptops to smaller screen sizes for mobile phones. Any questionnaire changes that are triggered by moving to online can bring differences in survey results.

Differences stemming from survey presentation can be most pronounced when moving from a telephone to an online survey because this means both a change from an aural mode of question delivery to a visual mode and a change from an interviewer-led to a self-administered approach. Differences can be less noticeable in a move from face-to-face to online because face-to-face often combines aural and visual presentation and interviewer-led and self-administered survey elements.

Aural vs. visual

Aural modes can encourage people to select the last option more often, because of a recency effect where people may more easily remember the last option read to them. Visual modes can encourage people to more often select the first option presented, owing to a primacy effect where people use the first answer option as a base of comparison for the other possible options.

For multi choice selections, an aural mode tends to force a yes/no for each answer item listed, whereas a visual mode can present the same question as a multi choice list requiring people to select only those items relevant (no response is required for "no"). As such, visual modes may encourage fewer answers. Aural modes offer the ability to hide "don't know" and "refused" options whereas these are more obvious in visual modes and may be selected more often.

Scale questions in visual modes have all response choices and labels presented fully to the respondent. In aural modes, scales are more likely to include only the end-point scale labels because it is tedious for interviewers to read out every scale label in full. This can lead to a tendency for aural modes to produce more extreme positive or negative responses at scale questions.

Interviewer effect

The presence of an interviewer can impact the respondents' candour. Agreeing in a social interaction is often easier than disagreeing and this can translate into more "agree" options on an agree/disagree scale in interviewer-administered surveys.

Social desirability bias is present in interviewer-administered modes and can lead to socially desirable responses. The

presence of an interviewer has greatest influence on attitudinal and ratings questions, less so on factual, neutral, or more objective questions. That said, some topics are more prone to the interviewer effect than others. People may be reluctant to disclose socially undesirable traits when discussing sensitive topics such as politics and even seemingly less sensitive topics can be affected; for example respondents may overplay frequency of tooth brushing in a survey on dental care administered by an interviewer.

In interviewer-administered modes the interviewer can motivate the respondent to complete each survey question and can probe, clarify and instruct respondents.

For example, unaided awareness can sometimes be much higher in face-to-face or CATI interviewing than online because the interviewer can continue to ask "any others?"

This may result in fewer cases of item non-responses or of unengagement than in self-administered modes. Self-completion surveys offer the respondent privacy to answer sensitive questions and levels of non-response for sensitive topics can be lower.

Figure 1 Survey presentation differences summary

Offline	Online
Last answer selected more frequently (telephone) / First answer selected more frequently (face-to-face)	First answer selected more frequently
More selections at multi-choice questions	Fewer selections at multi-choice questions
Don't know/refused options chosen less frequently	Don't know/refused options chosen more frequently
More extreme responses at scale questions	Fewer extreme responses at scale questions
More "agree" answers at scale questions	Fewer "agree" answers at scale questions
Fewer non-responses/richer responses to open questions (exception: sensitive topics)	More non-responses/ fewer rich responses to open questions (exception: sensitive topics)

SURVEY PRESENTATION

DO THE DIFFERENCES MATTER?

On one hand, the differences do matter. Offline and online surveys produce different results in terms of raw numbers. To give an example, people in an offline survey are likely to show different levels of brand awareness than people in an online survey. The size and direction of these differences will depend on the topic of research, country of field and target sample. They can be substantial. In research-on-research projects completed by Ipsos Interactive Services examining the differences between offline and online surveys in APAC region, between a third and a half of all survey data points showed significant differences between the offline and online answers[1].

Figure 2 Survey presentation differences summary

Indonesia
- No significant differences
- Significant differences

49% / 51%

Thailand
- No significant differences
- Significant differences

36% / 64%

India
- No significant differences
- Significant differences

51% / 49%
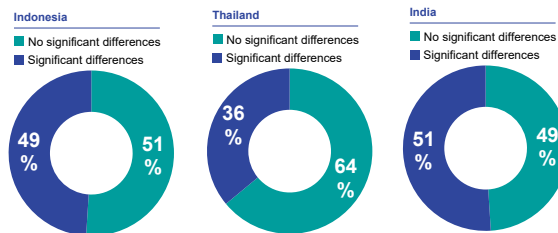
Source: [1]Research-on-research completed in APAC, 2018-2019

As such, online and offline are not interchangeable. Regardless of market, in general, if the same project was to be run offline and online, they would produce different results. This means that a project run offline cannot be run online and maintain data consistency.

On the other hand, these same differences may not matter. Extensive parallel testing experience at Ipsos has shown that research insights and outcomes are often correlated between offline and online results.

To use to our earlier example, brand awareness might register different levels in an offline and an online survey but if the data is highly correlated the awareness of the brands relative to each other is the same. If brand A scores the

highest in offline, brand A also scores the highest in online. In such circumstances the transition to online would not bring disruption to brand management, the relative results and conclusions drawn would be similar.

In extensive parallel testing coordinated by Ipsos Interactive Services across 19 studies in Latin America, high correlations were observed in the business decisions produced by the offline and online surveys across a variety of key performance metrics[2].

These results echoed those of earlier parallel testing programs completed by the Innovation service line evaluating concept testing in North America and Europe. At this time, a commonly used offline data collection method was mail surveys and as such online was tested alongside mail in this parallel program. 42 packaged goods concepts were
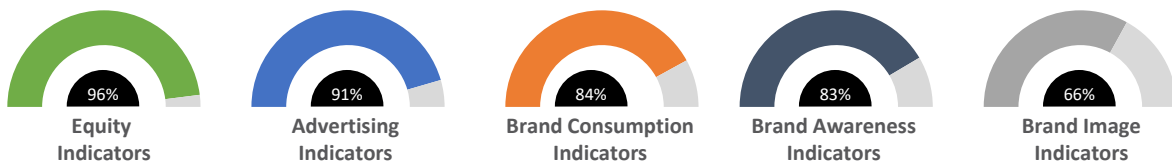
tested across four key performance indicators and the high correlations meant that the resulting business decisions were similar across the two methods[3].

Figure 4 Survey presentation differences summary

Metric	Correlation
Top box purchase intent	0.80
Top 2 box purchase intent	0.84
Value	0.93
Uniqueness	0.94

Source: [3]Parallel testing in North America and Europe, 2000-2004

Figure 3 Survey presentation differences summary

96%	91%	84%	83%	66%
Equity Indicators	Advertising Indicators	Brand Consumption Indicators	Brand Awareness Indicators	Brand Image Indicators

Source: [2]Parallel testing in Latin America, 2012-2015

HOW SHOULD I MANAGE THE DIFFERENCES?

Every research program is different which is why we assess transitioning on a case by case basis. Where online samples do not provide appropriate representation of the consumer population of interest, a transition is not recommended. Where online is a good fit, transitions can be managed by either making a clean break with the past or by measuring data trend changes through parallel testing. Where data trend changes are correlated, data calibration (if necessary) can be applied to re-state the historical data to be directly comparable with the new online data on key measures.

Clean break

For programs where multiple study elements are changing with the move to online, it is often not practical to maintain data trends. Survey re-design elements include questionnaire content, sample definition and structure and fieldwork timings. With many things changing at once, it is not logical to track the same trends. With a clean break, transitions can be swift and research budgets are not diverted to parallel testing.

Parallel testing

Data trend changes can be measured with parallel testing where program elements can be kept consistent. In parallel testing the study is run offline and online at the same time, to isolate the effects of the online transition from changing market conditions. It allows for relationships between the offline and online data to be measured through data correlations, so that the online results can be understood in the context of the offline results.

Where data correlations are high, the differences in trend results are systematic. In our brand awareness example, the rank order of the brands would be the same (or similar) in offline and online, and the online results would lead to the

same business decisions. Certain research markets, categories and brands lend themselves more than others to high correlations and as such are easier to transition to online than others. Research categories that are relevant to technology often produce more differences, for example an online bank may be ranked differently in a financial services survey by offline and online consumers.

The following data is illustrative of good correlations between

1. mail and online packaged goods purchase intent scores completed in North America by our Innovation service line[4]
2. face to face and online brand awareness for haircare brands in APAC[5]

Data calibration

Differences can, if necessary, be reconciled through data calibration in which the offline data is adjusted to be directly comparable with the online data. Data calibration is applied in situations where the online trend data needs to be viewed side by side with historical trend data.

Different metrics react to a transition to online with different intensity meaning that some trends will likely change more than others with a move to online. The calculations required are performed measure by measure because the same calibration factor cannot be applied to all.

DATA CORRELATION

The correlation between offline and online survey answers is measured using a correlation coefficient. Correlation coefficient scores range between -1.0 and +1.0. +1.0 is a perfect correlation score. Scores above +0.7 are generally considered as good correlations indicating that the results lead to the same business decisions. Scores of +0.5 to +0.7 indicate reasonable positive correlations and 0 to +0.5 shows weaker or no correlation. A negative value of the correlation coefficient score indicates an inverse association where the ranking of answers would be reversed.

Figure 5 Online vs. Mail Parallel Test Results For Packaged Goods Concepts
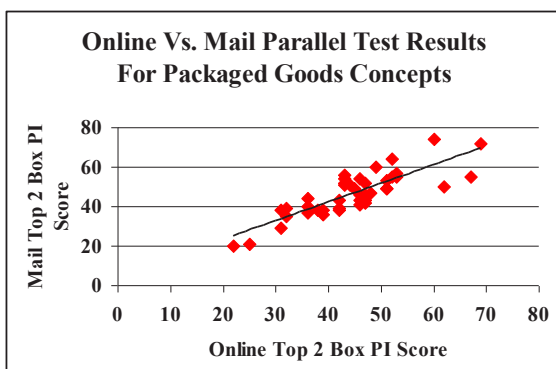
Source: [4]Parallel testing in US, 2000

Figure 6 Brand awareness
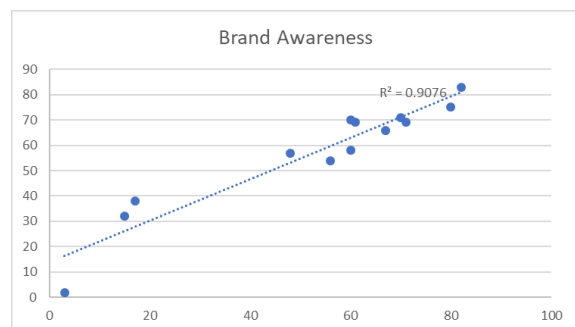
Source: [5]Research-on-research in India, 2019

IS THERE A WAY OF MINIMISING THE DIFFERENCES?

Steps can be taken to minimise the results differences that moving to online will likely bring, though it is important to recognise that it is impossible for all measures to match perfectly or to show high correlations between the offline and online results.

MATCHING SURVEY PRESENTATION

Keep survey parameters consistent wherever it is practical to do so. Retain the same questions within the survey in the same order, though it is likely that some modifications will be required to fit an online approach. Minimise differences coming from survey presentation changes by rotating or randomising the order in which survey answer options are presented to mitigate potential primacy and recency effects.

MATCHING SURVEY POPULATIONS

For programs where the same survey audience can be reached offline and online, match sample structures on key demographic characteristics (such as gender, age, region, socio-economic status). Even where sample structures can be matched however, there can still be differences in the profiles of non-quota demographics offline and online.

In our recent parallel testing in APAC region, we included broad controls for socio-economic groupings, but nevertheless, people reached online skewed towards the top end of the socio-economic classifications. In India, for example, the proportions of NCCS A / NCCS B

were controlled by sample quota, but within these broad classifications there was a skew in online within NCCS A towards A1, the most affluent part of the NCCS A groupings[6].

	A1	A2	A3	B1	B2
Offline	20	32	30	13	5
Online	55	19	6	14	6
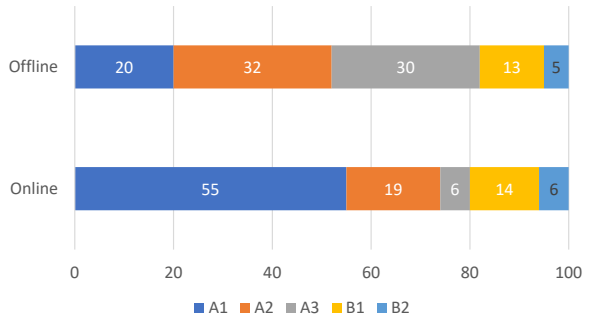
Source: [6]Research-on-research in India, 2019

Chi-squared automatic interaction detector (CHAID) analysis can be used to review the heterogeneity of key metrics results amongst each demographic subset to identify those parts of the sample that are most influencing results differences. This analysis is used to inform sample structure decisions towards minimising differences amongst key metric data trends when transitioning research programs.

Essentially, datasets are investigated with cross analysis of demographic sub-groups to understand what is driving results differences and whether there is a consistent pattern. Where cons6stent patterns are detected by sub-group, online sample structures are designed to control for those demographics that have the most bearing on key metrics outcomes, using fieldwork quota and/or post field data weighting.

BEST PRACTICES FOR TRANSITIONING FROM OFFLINE TO ONLINE

Anticipate change

Methodology choice has a bearing on research findings, driven by differences in population and survey presentation. The raw numbers generated by the same survey run offline and online are likely to be substantially different. The results though different, can be highly correlated, and in these situations the differences do not disrupt business management, the decisions taken following the research program results are similar.

Make a clean break

Transition studies to online at a natural break point in the program (for example as part of a larger survey re-design involving elements such as questionnaire changes, target sample changes). A clean break can be made with historical data with no requirement to manage data trends.

Parallel test

Measure the differences specific to a given set of study parameters with parallel testing. The new results generated online can be understood in the context of the offline results. For tracking programs a two or three month parallel is ideal, because a one month parallel could be adversely impacted by unusual market conditions. For data comparisons, we recommend to work with aggregated data collected over the parallel waves to even out any shorter data set fluctuations.

Measure correlations

Data correlations are run on key survey metrics to determine if the differences between the offline and online results are systematic. Highly correlated data indicates that though the online answer is different from the offline answer, the relativities between the answers in each dataset are similar and would lead to similar business decisions.

Calibrate data

Apply statistical recalibration in situations where data changes are correlated and where there is a need to view the online results side by side with the offline results. Through data calibration, the offline results are adjusted into the new online context. Each measure requires calibrating individually and it is practical to calibrate key measures only. Key measures include awareness, usage, equity, image, purchasing and satisfaction.

Minimise differences

Where there is a requirement as part of a transition to view new data in the context of the old, match survey parameters to be consistent with the offline version, as much as the online method permits. Ask the same survey questions in the same order to a matched consumer sample using the new online data collection channel.

Analyse sample sub-groups to identify sample profile characteristics that have the most bearing on key metrics results differences. Once identified, design online sample structures with these in mind to mitigate differences and encourage higher correlations with offline results. It might be that simply replicating the sample structure in online does not produce highly correlated results. Quota controls may be needed on new variables to better safeguard data consistency.

Online research continues to evolve and offer new capabilities. Access to consumers and citizens continues to grow such that Ipsos will be able to reach upwards of a 100 million people. It may be that the decision to move from offline to online research should not be "if", but when.

Irrespective of these general rules, each programme should be considered on a case-by-case basis. Please reach out to your local Ipsos researcher to discuss the plan for your research programme.

REFERENCES

1. Research-on-research completed in APAC, 2018-2019

2. Parallel testing in Latin America, 2012-2015

3. Parallel testing in North America and Europe, 2000-2004

4. Parallel testing in US, 2000

5. Research-on-research in India, 2019

6. Research-on-research in India, 2019

A NEW WORLD

Sara Wilkinson Director, IIS Global Research on Research

Leah McTiernan Global Director, Business and Community Development, Total Operations

www.ipsos.com
@Ipsos

GAME CHANGERS