



## Are my marketing campaigns having an effect – **is it the desired effect?**



This is a key question considered by most direct marketers and asked of direct marketers by their senior management. For some reason, however, the ability and time invested by some organisations to ‘estimate’ the performance of their chosen campaigns, in the form of business cases in advance, is greater than the investment in empirically evaluating the actual impact of campaigns after the fact.

### **A quick CRM (BTL Marketing) strategy check**

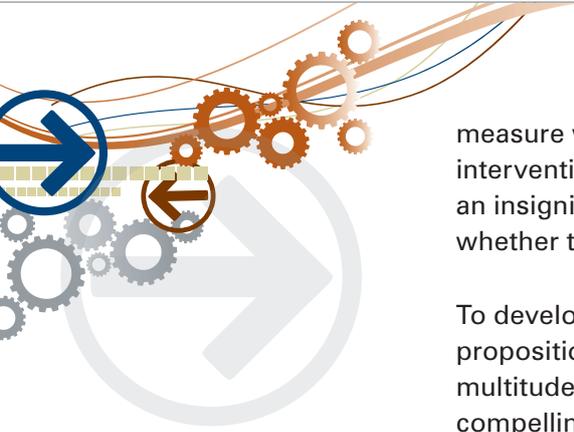
Why is CRM capability deployed and why are below the line (BTL) marketing campaigns executed..? Generically it is because there is a belief that when highly targeted groups of customers are presented with relevant, timely and compelling offers they will ACT. These almost always look to create an underlying change in behaviour of some form and more often than not promise the business (and are given commercial approval) on the basis of incremental cash to the bottom line; through incremental revenues, cost savings or both.

The first step in evaluating your campaign effectiveness, and to establish causality between a campaign and a change in behaviour is to isolate, identify and select an appropriate (sizable and random, but representative) “control group” for each campaign activity. Alongside these campaign or campaign-cell level control groups it represents best practice to hold out a representative universal control that can be used to evaluate the end to end impact of the entire BTL marketing strategy.

### **How clients should define and put together effective and optimal control groups**

First, what is a “control group” and what is its purpose? The control group is an accurate representation of the customer group targeted by a campaign and provides the baseline ‘do nothing’ reference against which the customers that are targeted with the campaign can be compared. A statistical significance

## Are my marketing campaigns having an effect – **Is it the desired effect?**



measure will prove beyond reasonable doubt whether the chosen campaign intervention was causal and if it did elicit a significantly positive, negative or an insignificant impact on target customer group behaviour and importantly whether these results will be repeatable.

To develop a winning campaign portfolio especially in the transaction/proposition rich communications services market typically requires testing a multitude of permutations of highly targeted proposition, relevant, timely and compelling. To effectively compete or keep pace requires intuitive CRM genius, lots of trial and error (minimised through experience) or ideally both in equal measure. In the complex, dynamic and competitive CSP marketplace, even the greatest conviction marketers will only get things partially right a lot of the time.

In the data driven 'Test and Learn' phase of campaign development many campaign cells/tests will be evaluated. Some will prove to be an instant 'success' and worth deploying in operational scale. Some will be shown to have material detrimental effects, but expect that many more will not prove to be particularly significant or successful at the first iteration.

The ratio of material successes to total activities during campaign development and optimisation dictates that organisations have some form of automated first stage evaluation as it is highly unlikely, and indeed inefficient, to have sufficient resource to evaluate each and every campaign manually.

Potentially as many as 4 out of 5 campaigns will be shown to have no material/significant effect in the initial test and learn phase of development, or be shown to be no more beneficial than the 'Champion' during 'Champion/Challenger'<sup>1</sup> optimisation. This can be down to the combination of call to action and reward simply not being relevant and compelling enough to elicit the desired responses. However, on (many) occasions, especially during test and learn, the ability to measure a significantly positive or negative impact is simply a result of poor campaign design and/or execution. As a consequence the material changes occurring in the small subset of the target group that do respond positively are drowned out by the sheer volume of the target that simply didn't receive the intended message, didn't read it when delivered, didn't consider it relevant or compelling.

---

<sup>1</sup> Champion/Challenger refers to the practice of making operational a campaign that is known to have a significant positive impact whilst continuing to run small but statistically significant tests of variations of the campaign to improve on the uplift and RoI.

## Are my marketing campaigns having an effect – **Is it the desired effect?**



Developing a winning campaign portfolio is in significant part down to deploying a meticulous and iterative test and learn methodology. In the situation that significant campaign impact is not demonstrated by the first campaign execution some marketers, especially those who have opted for operational scale deployment and promised big benefits from day one, lose patience and focus too quickly and can resort (deliberately or otherwise) to claiming false campaign benefits by using convenient pseudo-science (eg. by comparing just a campaign responder group against the entire control). Such results can (and usually do) look very positive (and even robust to the untrained eye), but let us note that these analyses may represent a very distorted and misleading 'evaluation' that holds no evidence of the causality of the campaign. A campaign evaluation education piece is clearly necessary for such clients/organisations. Once the fundamentals of campaign evaluation are appreciated and the importance of them understood any deviation would be seen to amount to false justification with potentially highly damaging consequences.

### **The ultimate test and learn campaign development program (Mighty oaks from little acorns grow)**

Maximising the broadcast message before empirical evaluation is poor CRM practice. In many cases good, bad and indifferent conviction CRM tests (pseudo-experiments) are conducted in operational scale with a lack of attention paid to evaluation. Marketers pursuing this strategy are very bold and implicitly asserting that they know the outcome before they see real results. Worse still the execution of operational scale trials (which often end up being only loosely targeted) frequently yield immeasurable/insignificant results when appropriately evaluated.

Where the luxury of time to market is available the ultimate in test and learn campaign development is to run pilot campaigns. These keep the target and control cells of equal volume (50% to receive the campaign, 50% to be held out as control group) and ensure that these are of sufficiently size to ensure statistical significance of the expected response. Taking such an approach has significant advantages: it avoids running detrimental campaigns in operational scale, it minimises the depletion of the target customer pool that is required for further iteration/refinement, it minimises signalling eg to the competitor intelligence programs of competitors. One drawback is that potentially the benefits are delayed, however, this is in many/most instances a cost worth carrying.

## Are my marketing campaigns having an effect – **Is it the desired effect?**



### **Don't allow the low/no cost of SMS to cheapen/proliferate your marketing communications**

Especially to mobile operators, sending marketing messages via SMS may be relatively cheap versus other channels, and in some cases it is viewed as zero incremental cost. This should not be used as a rationale for sending out blanket campaigns with poor relevance/appeal (spamming). To do so will only serve to desensitise your customer base to your marketing and reduce the effectiveness of relevant marketing communications. It also, rather boldly and naively, assumes that your campaigns only can have positive and not negative effects. Note also that alongside your best in class targeted communication strategy increasingly other 3rd parties are looking to contact your customer base with text and other handset based campaigns. Where others employ tactics that are less considered and targeted than your own they might also have a negative impact on your ability to gain cut-through with SMS if you do not have a history of relevance.

### **Defining a basic control group**

There is not a great deal of science required to take a representative control group. Here are the steps:

- 1) Select the entire group of people that fit the intended profile for the campaign. The size of the entire group can have a physical maximum due to the number and nature of the parameters of the profile, but in many instances there will be a choice to select a small overall group for a pilot test and learn.
- 2) Use a random function to take out a sample (your Control) usually based on MSISDN. It is key to ensure that this sample is sufficiently large to enable you to measure with significance any changes in behaviour that arises as a result of the campaign. If the sample taken is large enough (a good value for decent 'security' to evaluate statistical significance would be  $\geq 2,000$ ) then the control group will be representative of the target at first attempt most of the time. Problems with equivalence become worse the smaller the control group and to a lesser extent the percentage the control group represents of the overall campaign sample and resampling.
- 3) Always perform a basic profile comparison exercise of the resultant campaign target vs the control group. Campaigns (assuming evaluation is required) should never be executed before this check is complete or

## Are my marketing campaigns having an effect – **Is it the desired effect?**



you run the danger of executing a campaign that is then impossible to evaluate. The minimum profile requires the measurement and tracking of the campaign response KPI plus:

- a. ARPU and/or AMPU
- b. a measure of unique 'active' customers
- c. a distribution curve of customers by tenure

(See Appendix 1 for an expansion of these points)

The purpose of these, and where appropriate further checks to be conducted before campaign execution, is to confirm that the control and target populations are indistinguishable from one another.

**Note:** it is these very same parameters that will make up the first stage of automated campaign evaluation; post campaign deviation of measures is sought where in the pre-campaign period no deviation is the requirement.

- 4) If the above checks result in target and control showing equivalence the go-ahead can be given for campaign execution. If they are not a new control group must be sampled (see Appendix 2 for notes on control group resampling).

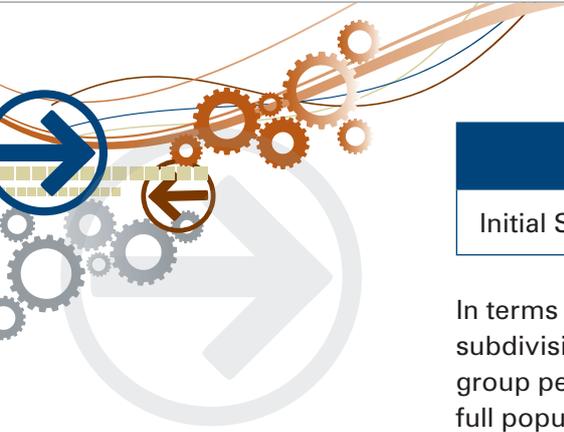
Having reached this stage you have a target group and a representative control and it is only now possible to kick off a valid BTL campaign test. However, attaining a representative control group is only the start of things, but this alone does not constitute a good marketing campaign test.

### The key to successful evaluation

Having established that a 'proper' control group (validated ahead of campaign execution) is the core requirement for successful evaluation, we will specifically critique why a campaign evaluation that compares only the campaign responders with the total control group leads to drawing invalid and incomplete conclusions (usually falsely positive). As a minimum set of KPIs we recommend tracking response rate, active customers, ARPU and total revenue.

Let us take a hypothetical example campaign which eg. aims to increase customers usage by offering a material reward. Firstly, a robust campaign size is selected and a random, and representative control group (in this case 50%) is taken (and its representative nature validated by pre-campaign analysis).

## Are my marketing campaigns having an effect – **Is it the desired effect?**



	Target Population	Control
Initial Selection	2,000	2,000

In terms of robust evaluation there is no need to worry about any further subdivisions of these populations. All hard evaluation of target versus control group performance must be made by comparing any/all KPIs across these two full populations.

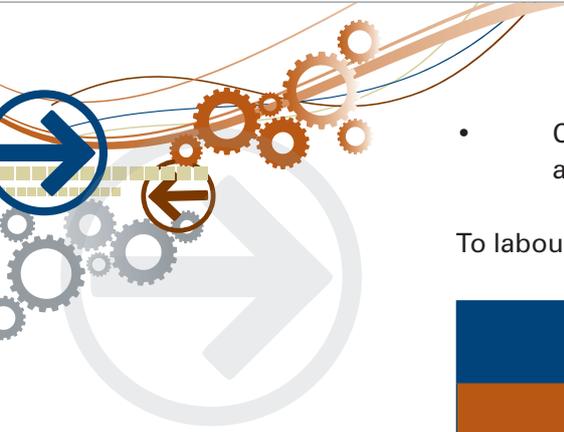
Let's assume in this campaign that 200 of the 2,000 'Target' customers became responders. Very poor practice then compares the attributes of just these 200 customers versus the total 2,000 Control group. Doing so typically introduces a high degree of auto/self-selection. This is the equivalent of comparing fresh apples with concentrated apple juice and there are two fundamental issues with this approach:

1. The actual running of a campaign restricts the population of customers who can possibly be influenced by the campaign. Tactics include message delivery failures, message deletion before opening/reading messages dismissed as irrelevant. This means that total target/control is not representative of sub segments of the target group.
2. In this hypothetical example if I assert that just 100 of the 200 gross responders were incremental responders (ie. uplift brought about directly by the campaign), then 100 responses were 'organic' (simply down to the course of natural events). This will also be evident by 100 responders in the control population despite these customers receiving no marketing message. Clearly this typical definition of campaign response (to increase in usage) naturally filters out a substantially higher ARPU subset of the entire population.

Looking at the gross responder group versus total control mis-reports (inflates) the campaign benefit by claiming credit for the underlying behavioural changes and self-selection in addition to any genuine change brought about as a direct consequence of the campaign. Very importantly, it also misses two other genuine and potentially significant changes brought about in customer behaviour. These are:

- Customers that did change behaviour, but just not significantly enough to be captured as a 'responder' – omitted positive effect

## Are my marketing campaigns having an effect – **Is it the desired effect?**



- Customers that did change behaviour, but responded to the campaign in a negative way

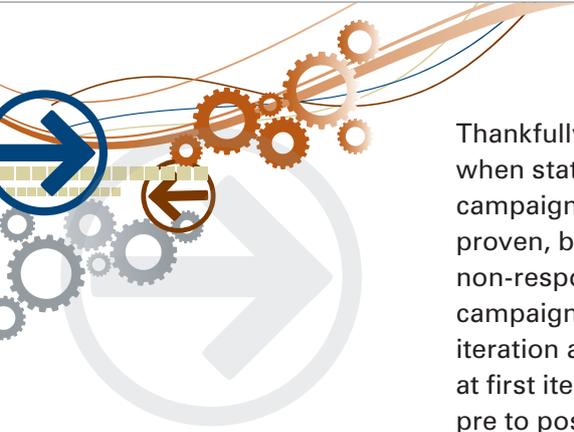
To labour this point a little further consider the table below:

	Campaign Customers		Responder Measurement		
	Target	Control	Target Response (Organic + Uplift)	Control Response (Organic)	Uplift/ Incremental Response (Target - Control)
Target Selection	2,000	2,000	200	100	100
Messages Delivered	1,400	n/a	180	n/a	?
Messages Opened/ Read	700	n/a	140	n/a	?
<i>Messages considered relevant</i>	<i>350</i>	<i>n/a</i>	<i>120</i>	<i>n/a</i>	<i>?</i>

**Note:** Shaded areas represent unknowns in most campaign situations.

The 200 gross responders from the campaign are the result of campaigning to the whole 2,000 Target. Here we confidently know we have a representative 2,000 control group. We can see that the gross response to the campaign target is 200 or 10%, but tracking the control group and observing 100 or 5% responses allows us to calculate the all important campaign impact; here an incremental 100 or 5% responders. As marketers we know that our marketing message can only have an effect when our message is delivered, when it is read (and when it is deemed relevant). The temptation can be strong to try to determine this very specific effect (message effect where it applies to just receptive targets), but note from the above there is no way to measure this like for like; and certainly not by comparing 200 responders versus the total control group. Even if such a valuation were attained it's value is entirely academic (in a hypothetical world where targeting miraculously became 100% accurate), so focus your energies and analytics on improving the targeting and improved evaluation will follow.

## Are my marketing campaigns having an effect – **Is it the desired effect?**



Thankfully for the deployment of BTL campaign capability all is far from lost when statistically significant uplift is not proven in the first execution of a campaign. There are many cases where robust evaluation of uplift is not proven, but valuable insights are gained. The results of campaign responders, non-responders and the control group can be compared in the pre and post campaign timeframes in order to focus in on campaigns worthy of subsequent iteration and more refined targeting. Some campaigns that were unsuccessful at first iteration will yield information about their potential benefits through the pre to post campaign KPIs of the responder group materially diverging from that of the control group (whilst the non-responder group follows a behaviour that mimics the control group). In this situation the ability to identify what differentiates the responders from the non-responders can lead to a refined campaign selection and a winning campaign from a subsequent iteration. Again the ability to filter and ascertain where the bespoke analysis and next round of campaign effort should be deployed should be guided by the automated campaign reporting.

### 5 Simple questions to help you build a portfolio of successful campaigns

1. Is my campaign highly targeted, relevant, timely and compelling ?
2. How will I measure success - define in advance what your key KPI(s) are – if necessary track these in addition to revenue, active customer count, response and reward KPIs all of which should be mandatory (and defaults in InTelestage Campaign Reporter)?
3. Do I have a representative control group - put approval processes in place that requires a 'sign off' of the control group before proceeding to campaign execution?
4. Show me the incremental change (uplift) measured in the Total Target Group versus the Total Control group. Never rely/evaluate based on gross response rates or responders only which will be very misleading in isolation?
5. Show me the behaviours of target responders versus none responders versus control group through a period covering pre and post campaign – ask how could the campaign have been more highly targeted, relevant, timely and compelling?

## Are my marketing campaigns having an effect – **Is it the desired effect?**



### In Summary

#### **Don't just guess and hope, better to test and learn!**

The starting point for BTL marketing is hypotheses about how the behaviour of targeted customer segments can be changed for business gain through sending relevant offers and communications. Being successful at BTL marketing is highly dependent on the number of pilot campaigns a company can run, and evaluate versus representative control groups, at a given time.

#### **Bryan Eastwood**

Bryan is an experienced business leader and base management expert with over 12 years of commercial experience in marketing, strategy and finance functions within the CSP market.

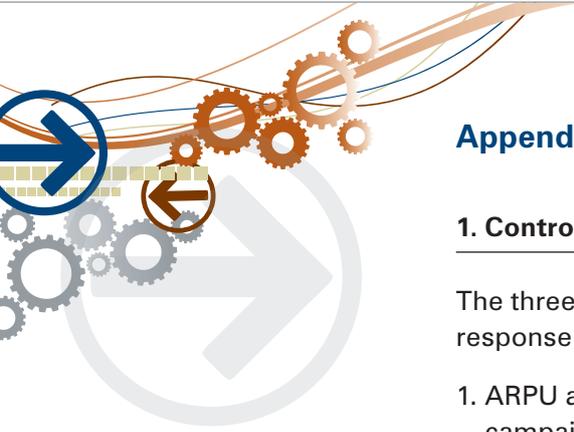
He has a track record of selling compelling actionable recommendations to CxO level management by taming, structuring and interpreting customer and financial data to create powerful commercial insights that add value to the bottom line.

Bryan was a senior marketing manager at Vodafone and T-Mobile prior to founding BOA Consulting Limited where Business Logic Systems are one of his key clients.

Business Logic Systems  
The Didcot Enterprise Centre  
Hawksworth  
Didcot  
Oxfordshire OX11 7PH

**T** +44 (0)1235 510427  
**E** [info@businesslogic.co.uk](mailto:info@businesslogic.co.uk)  
**W** [www.businesslogic.co.uk](http://www.businesslogic.co.uk)

## Are my marketing campaigns having an effect – **Is it the desired effect?**



### Appendices

#### 1. Control group profiling in more detail

---

The three default KPIs to use in campaign tracking, in addition to the campaign response KPI are:

1. ARPU and/or AMPU of both groups tracked for the periods up to the point of campaign selection
2. A measure of unique 'active' customers based on a pre-determined definition. In the absence of having an accepted active definition use any within the last 28day period. This period may be reduced materially also dependent on the nature of the customer profile and campaign call to action and reward. (Qn: could the definition of 'active' customer be defined by an input parameter into the InTelestage solution..?)
3. A distribution curve of customers by tenure (in the case of a postpay campaign this would be a distribution curve of customers by commitment point in/out of an contractual obligation)

There are many other possible segmentations that could be brought in to learn more about the differentiators between campaign responders and non-responders, but these are typically either already baked into your target and control group selection, or investigated through a piece of bespoke analysis to better refine a campaign that has had material effect or insignificant campaigns who's responders show promising behavioural inflections at the point of campaign.

#### 2. How to resample control group:

---

1. Simply take another random sample – eliminates a random aberration
2. Up the size of control (and target if necessary) and then take another random sample – should result in a representative control
3. If after a few attempts and an increase in sample size a representative sample is not achieved this probably indicates there is some systematic bias in either the random function or in the underlying data (eg. If a 'random' 10% sample were taken by selecting all MSISDNs that ended with '1', but all available numbers ending in 1 where selected to connect the initial customer, then this would clear not result in the desired random outcome). In this case you may need to find an alternative algorithm to take the random sample and then repeat.