Measuring Not Counting

Evaluating Social Marketing Communications

#IPASocialWorks

Ray Poynter, Fran Cassidy, & Simeon Duckworth

(with assistance from the many people listed in the Appendix)

Contents

Foreword

Social media marketing has come through the usual hype cycle and is now well established within the marketing armoury, although it's still developing and its role is still emerging. Or, rather, roles: like other digital media, social can be used in many different ways under the broad headings of marketing communications, customer insights, and customer experience and relationship management. It will take time for marketers to learn how best to exploit it under each of these headings.

This Guide focuses on the first category - the use of social media in marketing communications - although it also touches on the others. (Categories and distinctions in marketing are rarely clear-cut). The potential is obvious: social media have achieved huge global penetration and usage, and as marketers we naturally like to reach customers and prospects via all the media they use, provided we can find the right ways of doing so.

The potential is even greater today than five years ago, not only because of increased penetration and usage but also because of the parallel growth of the mobile internet: we can now reach consumers whenever we want and wherever they are and we also increasingly know where they are, when, as well as who else they talk to, when, and about what.

All this raises some big challenges. As with all digital technology, there are the usual issues of identity assurance and protecting customers' privacy and security, and of sorting and analysing the vast amount of data being generated. But social media marketing also raises some subtler behavioural challenges to do with finding messages and a tone of voice appropriate to the social media space. This is a private – or at least, semi-private – space. Consumers will appreciate brands and companies joining in their conversations only if they "fit in" and have something useful, interesting or entertaining to say. And, of course, it is the consumers themselves who judge what is useful, interesting or entertaining.

To exploit the potential and cope with these challenges, social media marketing is evolving. In the early days, techies and the more evangelical types of media folk saw it as completely revolutionary, separate from and probably soon replacing the rest of marketing, with brands joining people's private conversations - or enticing them as they walked through town – with exciting special offers and other messages. All this turned out to be not only more difficult technically than had been assumed (if you ask an IT guy "Can you do this?" and he answers "Yes", he's using that word in a special sense) but also often wildly inappropriate in terms of what consumers wanted and welcomed.

If used badly, social media, combined with mobile and location sensing, have a capacity for annoying customers surpassed only by rude sales and service people. In contrast, success involves combining the new channels and techniques with traditional capabilities like consumer understanding, creativity and measurement. This combination is at the heart of the current evolution of social media marketing.

Creativity is especially important. The best social media campaigns are those that go viral. This is the ultimate creative challenge: the message and its execution need to be so good (useful, interesting or entertaining – especially, funny) that large numbers of people pass it on to their friends and contacts purely for the social benefit of doing so. The potential ROI is huge, but largely depends on the quality of the creative idea and execution. The increasing recognition of this is part of the "growing up" process of social media marketing.

Another part of this evolution is to do with measuring success. During the late 1990s, the digerati put a lot of emphasis on eyeballs. There was a reluctance to focus on old-fashioned things like revenue. When the bubble burst in spring 2000, there was a brief period when digital marketing was no longer seen as a hot topic. But when it came to resource allocation, marketers were still increasing their digital investment year by year, as they have done almost relentlessly ever since

(Barwise & Farley, 2005). At the same time, clients were no longer impressed by eyeballs: they wanted to know how many people clicked on an ad and, if possible, how many of them then went on to take up the offer, buy the product, or whatever the aim was.

Something similar is now happening with social media marketing – hence the title of this Guide. Clients need to know that money spent on social is delivering value. This means measuring impacts and outcomes, not just counting clicks, likes and interactions. Like the internet ten years ago, social media marketing is now coming of age.

Measuring social is hard because:

- It needs to work in combination with other channels and activities
- It is dynamic and interactive, which raises a number of measurement challenges
- Social media campaigns can have many different aims
- Campaign budgets tend to be smaller than those used in media such as TV although larger ones are appearing and measurement approaches need to reflect this diversity.

On the plus side, like all digital media, social can generate a lot of data at little or no incremental cost and lends itself to the test and learn approach traditionally associated with direct marketing - but cheaper, much, much faster and, usually, on a larger scale.

As the importance of social grows, along with the resources allocated to it, there is a growing need to evaluate social with the same rigour as that applied to more traditional channels, and increasingly with metrics that take an integrated, as opposed to channel-specific, approach.

This Guide is well timed to take stock as social media marketing communications come of age. It is the result of a collaboration between the IPA, MRS and Marketing Society, funded by the three major platforms: Facebook, Twitter and LinkedIn. It summarises the current state of play and offers guidelines on how to make the most of this young but rapidly maturing medium, based on and illustrated by some great case studies. I'm not aware of any other publication anywhere that addresses these issues so comprehensively. I think it will prove extremely useful and provide a foundation for the rapid further development of social media marketing.

Patrick Barwise
London Business School
September 2014

Introduction

Social is still in its infancy and rapid innovation has led to a proliferation of data, methods, and language. Nobody can claim to understand fully how social works in the marketplace and anyway its nature and use are still evolving. However, some key points are becoming clearer, such as the need to measure the impact of social activities on relevant metrics and the limited value of simple concepts such as counting the numbers of posts, followers, and likes.

The lack of established practice on the evaluation and measurement of social is costly and is holding back innovation. It is costly because people are not sure of the best ways to optimise their use of social. It is holding back innovation and slowing down adoption because organisations are unwilling to invest more unless they can evaluate the benefits more accurately. Therefore, a cross-industry initiative called #IPASOCIALWORKS has been formed to help highlight good practice and advice on measuring and evaluating social.

#IPASOCIALWORKS brings together the Institute of Practitioners in Advertising (IPA), the Market Research Society (MRS), the Marketing Society, and the key social platforms (Twitter, Facebook, & LinkedIn). This is believed to be the first time that such a wide collaboration across the marketing and advertising industry has been formed.

This project is ongoing and its key deliverable is this 'Guide to Measuring Not Counting', a substantial body of current practice and learning. In addition there are summaries, short notes, links to case studies, and there will be a series of events.

Key Assumptions

The project is shaped and guided by the following assumptions:

- Social should be held to the same sort of standard as other forms of media and activity. If social is held to a higher standard, opportunities will be missed; if it is held to a lower standard then it will tend to be considered as a fringe option.
- Social is not just a medium and it use is much wider than marketing. In order to make the project practicable it has been focused on marketing related issues and specifically on the evaluation of marketing campaigns and activities.
- Social does not always fit the traditional model of a campaign, with a clear start and finish. In social it is necessary to think of campaigns and activities comprising a variety of forms and timespans.
- The measurement of campaigns and activities in social should build on what is already known about evaluating campaigns and activities in other media.
- Organisations spending money need to be able to identify the value they are achieving, so social needs to be measured/evaluated in terms of metrics that relate to business objectives.

It is recognised that there are many interesting uses and applications of social that are not included within this project. Other uses include: using social to create marketing ideas; as a source of research about things like customer satisfaction, product usage, and brand awareness; and the use of qualitative research in the context of social. However, these wider issues are largely outside the scope of this Guide. It is envisaged that there might be future projects, for example looking at social as a source of insight and social as a method of delivering services and interacting with customers/users.

Background to #IPASocialWorks

#IPASocialWorks is an industry-wide project to help advertisers, advertising agencies, researchers and social media owners develop a more robust approach to measuring the effectiveness of marketing through these new platforms. It is supported by the IPA in partnership with MRS and The Marketing Society, and sponsored by LinkedIn, Facebook and Twitter

#IPASOCIALWORKS has as its aim the creation of a comprehensive knowledge bank of social media effectiveness – the equivalent of the IPA's 30 years of peer-reviewed advertising effectiveness data set or the Marketing Society's Awards for Excellence.

The need for this project arises because of a perceived mismatch between the measurability of social and limited amount of evidence into how and when it works. A large part marketing's interest in deploying of social media is based the ability of these platforms to create data and be measurable in the broadest senses. However, many businesses struggle to see the connection between social data (for example a tweet) and outcomes (e.g. sales or profit). The advertising industry has spent many years addressing similar perceptions about its effectiveness in the context of more traditional media and many of its learnings need to be leveraged in this new context.

The relevant learnings from evaluating marketing communications in the context of traditional channels include:

- Rooting effectiveness in business metrics, as opposed to marketing metrics.

- Distinguishing between correlation and causation.

- 'Baking in' effectiveness thinking from the start.

One of the key challenges for utilising social is its lack of history and benchmarks. Traditional advertising media have developed reasonably slowly over the last couple of decades, allowing advertising and marketing systems to evolve with them. The social media ecosystem is relatively new and is continuing to change and evolve. While the old media world suffered from a paucity of data, the new social world is awash with data. However, that data is differing quality and standards. Whilst it took the UK advertising industry more than 30 years to create its current understanding, social media measurements need to be useful and reliable now.

The IPASOCIALWORKS Case Studies

A team led by Fran Cassidy has been reviewing a large number of case studies to identify those that show the effectiveness of social and which highlight the use of robust evaluation. The collection of case studies is growing all the time and can be accessed from the IPA website at: http://www.ipa.co.uk/page/IPA-Social-Works-cases.

This Guide makes several references to these cases, and to a wider set of cases which are outlined in the appendix.

Social and Marketing

This chapter clarifies the way that the term social is being used in this Guide, in the context of marketing campaigns and activities. The chapter sets out some definitions and looks at how social fits in the wider marketing ecosystem.

Defining and using social

Like many phenomena, social is hard to define, but easy to recognise. Facebook and Twitter are social. In contrast, in their traditional forms newspapers, billboards, and direct mail letters are not. The key features that are at the heart of social are:

- Interactive and social. Users of social can share, like, link, create, amend, and connect with the material and with other users, i.e. there needs to be a peer-to-peer element.
- Digital. Whilst there has always been social, in the context of this project the focus is digital.
- Attributable media. Not all social media is attributable, but social is accepted as a highly measurable form of media. Attributable media means it is often possible to determine who saw what, who interacted with what, and what they did next.

The evolution of social

When social media first burst on the scene, with networks such as MySpace and Friendster, it was a distinct sector and one that created ambiguity about how it would be financed in the medium and long term. Since then there have been three major developments:

1. The adoption of social by wide range of sites and services (for example news services, travel, and retail), seeking to get people to share, like, post, and co-create.
2. The creation of ad platforms that leverage their social members.
3. The movement towards mobile is shaping the way social is being used, particularly in the context of SoLoMo, the combination of social, location (e.g. GPS and beacons), and mobile.

Using social

There are a number of definitions of how social is defined and used. For example, Patrick Barwise, Emeritus Professor of Marketing at London Business School, likes to divide the marketing uses of social media into Marcomms (essentially, advertising), Customer service, and Customer Insight (Barwise & Meehan, 2010).

Similarly, McKinsey (Divol et al 2012) have defined four marketing related categories:

- **Monitor,** social channels for trends/insights. Monitoring implies researching what people are saying in social media (for example posts) and what they are doing (for example sharing, viewing, re-tweeting). This sort of activity is also referred to as social media mining, social media research, and listening research.

- **Respond,** to consumers' comments. This includes crisis management and the provision of customer services.

- **Amplify,** current positive activity/tone. Amplify is a broad term including: fostering communities, organising and promoting referrals and recommendations, and enhancing brand advocacy.

- **Lead,** e.g. create changes in sentiment and behaviour. This category includes a broad range of marcomms and promotional activities, such as awareness, launches, and deals.

This Guide focuses on the marcomms elements, consequently, issues such as service provision and crisis management are only tackled when they are closely linked marcomms related activities.

Social Media and the PEOM Framework

Social is often described in terms of the POEM model, which stands for Paid, Owned, and Earned Media (Corcoran, 2009). Each of the elements of the POEM model creates specific opportunities and challenges.

Paid Media

Paid media is the most traditional of the three types of media and refers to more than just social media. For example, paid media include:

- Print and TV advertising
- Radio and display advertising
- Direct mail
- Paid search and ad words
- Paid advertising in social networks

The Doritos campaign, highlighted in the case studies, centred on a paid media campaign. The paid media had as one of its aims driving people to owned media (i.e. its Facebook page). The Doritos campaign highlights the way that brands often seek to combine different elements of POEM.

Owned Media

Owned media means owned by the client, for example a corporate website, Facebook page, or Twitter account. Brands vary massively in terms of the scale of their owned media. Some brands having access to very large communities on platforms like Facebook, Twitter, and LinkedIn. For example, in January 2014, Coca-Cola had about 2.5 million Facebook fans from the UK and nearly 80 million globally.

Owned media includes other media owned by the client, including campaign sites, competitions, apps, as well as older forms of owned media such magazines and newsletters. Although owned media is sometimes referred to as free, it still incurs the costs of time and effort, the materials shared via owned media cost money too, and there is the substantial capital cost of having owned media.

Earned Media

Earned media refers to material that is shared or distributed by people who are not being paid for what they are doing. In social media this term includes review sites, shares, uploads, replies, re-tweets, 'favouriting', discussions in forums and on bulletin boards. It is sometimes referred to as C2C – customer to customer.

The 2013 Nielsen Global Trust in Advertising and Brand Messages report ranked earned media as the most trusted, followed by owned media, with paid media being the least trusted of the three. Interestingly, the trust levels reported for Europe are much lower for almost everything, compared with other regions.

Authenticity of Earned Media

One concern about earned media is whether it is genuine. There have been high profile cases where online reviews or blog posts that appeared to be from fans or neutral third-parties turned out to be paid placements. For example, author Stephen Leather has confessed to creating fake reviews on Amazon to help create buzz around his books (Charman, 2012).

Converged Media

Converged media is where two or more elements of the POEM model are combined. This tends to bring content to the fore rather than just volume (DerMers, 2013). The idea underpinning a converged media strategy is that if material is sufficiently interesting people will want to share it, link

to it, add comments etc. The paid or owned media is used to provide the resources for the earned media to utilise, and to generate sufficient interest to cause people to create earned media.

Social Marketing Approaches

Whilst most traditional forms of marketing are amenable to being utilised via social, eWOM (e-Word of Mouth) and COBRA (Consumers Online Brand Related Activities) are more specifically associate with social.

eWOM

Word of mouth (WOM) is as old as marketing and has long been considered the most powerful element in marketing, and the digital equivalent is eWOM. A simple example of eWOM might be a customer review section on a hotel website site, where the hotel seeks to create positive flows of information and advocacy from visitors to the site.

WOMMA (the Word of Mouth Marketing Association) draw a distinction between organic and amplified WOM. Organic refers to naturally occurring patterns of posts, links, shares etc. Amplified refers to marketing campaigns created to utilise the mechanics of WOM. One key feature to of eWOM is that it can be positive or negative, in contrast to most other routes which are only occasionally negative.

McKinsey suggest dividing WOM into three categories (Bughin et al, 2010).

- Experiential, which they claim accounts for 50%-80% of WOM and results from people's direct experience with a brand or service. This usually refers to where the experience does not match expectations. If something is as good or as bad as expected then there is little to say. Something like lost airline baggage (or a broken Taylor guitar) is so powerful because it is not what is expected.

- Consequential. This describes what happens when people are exposed to marketing.

- Intentional. This describes paid WOM, for example celebrity endorsements or paid influencer campaigns.

COBRAs

COBRA stands for Consumers Online Brand Related Activities. A COBRA can be as simple as persuading customers to view a video, through to uploading a picture of their new shoes/hat/meal to Facebook or engaging in an ongoing activity such as Sharpie's series of social media campaigns to get people to create and upload doodles in a variety of forms to a variety of sites, using sharpies.

COBRA can also be thought of as comprising three elements: consuming, contributing, and creating (Mutinga et al, 2011)

1. Consuming, for example viewing videos, sending branded gifts, and playing branded games.
2. Contributing, for example engaging in brand related conversation, including:
   - Rating or scoring products, e.g. Trip advisor.
   - Engaging in branded conversation, for example on Facebook page.
   - Adding comments to a brand site.
3. Creating, including uploading brand related videos, photos, drawings etc, writing brand related articles, creating product suggestions and ideas, for example Lego or MyStarbucksIdea

The uses of social in marketing campaigns and activities

Social works, in marketing campaigns and activities, in a variety of ways including:

1. Telling deeper and richer stories
2. Leveraging social behaviour
3. Enhancing/amplifying other activities
4. Monitoring and responding to discourses
5. Allowing real-time management of campaigns

These five are explored more fully below. Definitions are rarely precise in marketing and there were a number of other routes that could have been added to these five, for example making product and service delivery social, but they fall outside the scope of this Guide.

1. Telling deeper and richer stories

Social facilitates deeper and richer stories in a number of ways, including:

- Providing more space/time for a story to be told in its entirety, rather than fitting it to a conventional format.
- Allowing the story to expand over time, in response to reactions, questions, suggestions.
- Utilising co-creation, where the story is made deeper and richer through the contributions of the crowd.

Two of the case studies reviewed by the IPA illustrate how deeper and richer stories can be used.

- Visit Iceland utilised social to encourage Icelanders and previous visitors to share their stores of Iceland, as part of a campaign to re-establish Iceland following the financial crisis and the Eyjafjallajokull volcano. The campaign created rich and deep stories and generated a return on the marketing investment of 61 to 1, on a campaign budget of £2.24 million. (See case study).

- The Mattessons brand Fridge Raiders campaign was able to utilise social to create an engaging and rich story, which helped to arrest a decline in sales. The brand teamed up with a celebrity game (with over 3 million connections) to use followers to crowd source a "hands free snacking device". The activity led to a 20% increase in sales and a ROI of £2.44 for every £1. (See case study).

2. Leveraging social behaviour

One of strengths of social is that it can utilise social behaviour, for example to spread a message or promote an activity. Two examples that illustrate this are the ASB Bank and the Miss Universe competition.

- **ASB Bank** is a New Zealand bank that created a product called the "Like Loan", where an offered interest rate was lowered in proportion to the number of people who visited the product's Facebook page – at the end of the promotion the period one lucky visitor was then offered the rate. The campaign is reported as costing NZD 378K and the value of conversations to date exceed NZD 4.5 million. (Case study in appendix)

- The **Miss Universe** competition attracts viewers and voters from about 160 countries, requiring 23 million votes in 9 minutes. NBC created Facebook and Twitter buttons on a second screen app, which allowed people to 'share' the message that they had just voted on with their social media connections. Analytics showed that the social sharing has led to an increase of 12% in second screen activity. (Case study in appendix)

3 Enhancing/amplifying other activities

Generally, the most effective way to use social is in conjunction with other activities, in particular to amplify the impact of other channels. The examples that illustrate the how social can amplify and enhance other activities are shown by campaigns from Doritos and the Philippines.

- **Doritos** decided to broaden their brand appeal in terms of situations and users and launched a campaign based on creating a 'Mariachi Doritos', a band playing cheesy covers of pop songs. Social was used to gain attention, to recruit participants, and to help filter material. Items that achieved a certain level of social success were then elevated to paid media, to broaden their impact, which in turn fed the social elements. Data from Nielsen suggest that both soft measures (such as purchase intent) and sales increased. (Case study in appendix).

- The **Philippines** used social behaviour to amplify its campaign "It's more fun in the Philippines". For example, an app was created that allowed locals and visitors to superimpose the slogan on a picture and to post it to social. The campaign is associated with an increase in Google traffic and an increase in tourists, and was the 2013 winner of Warc Prize for Asian Strategy. (Case study in appendix)

4. Monitoring and responding to discourses

The ability to listen to conversations taking place in social media creates a number of opportunities, including:

- Gathering a 360 degree picture of the brand, its campaigns, and its activities.
- Sourcing of information about the activities of other brands.
- Gathering inspiration and advice.
- Checking for problems, underperformance, and alerts.

One companies that provides a great example of turning listening and responding into a campaign is BT.

- BT have developed a monitoring tool called Debatescape to help them understand what people were saying about their service and the brand. Amongst other uses, Debatscape allows BT to find people who are having problems and offer help. Being more proactive was part of a broader campaign that has resulted in savings, improved satisfaction, and reduced churn. (See case study)

5. Allowing real-time management of campaigns

Historically, the measurement of campaigns and activities was considered separately to the implementation and management of the campaigns/activities. Typically, the measurement was a method of evaluating campaigns after they had finished. However, with social the measurement and evaluation can be integrated and the measurement process used to guide and influence the implementation.

For example, O2 has developed a platform called RTO2 which monitors social conversations about O2, its campaigns, and its competitors. This information is used to provide monitoring and social delivery, but it is also used to manage and instigate activities. The Social Media Response team use the feedback from RTO2 to guide their work. In July 2012 O2 suffered a network outage for 10 million customers over two days, and was able to use their team to respond via social to minimise the negative impact, responding to every mention that appeared in social media. (Case study in appendix).

The interaction of social with other media

In most cases, social is used in conjunction with other media, and in most cases its impact will be partly direct and partly through its interaction with other activities. This level of interaction needs to be designed into the campaign/activity at the outset and should be addressed by the evaluation.

The sort of interaction that should be expected is shown in the diagram below.
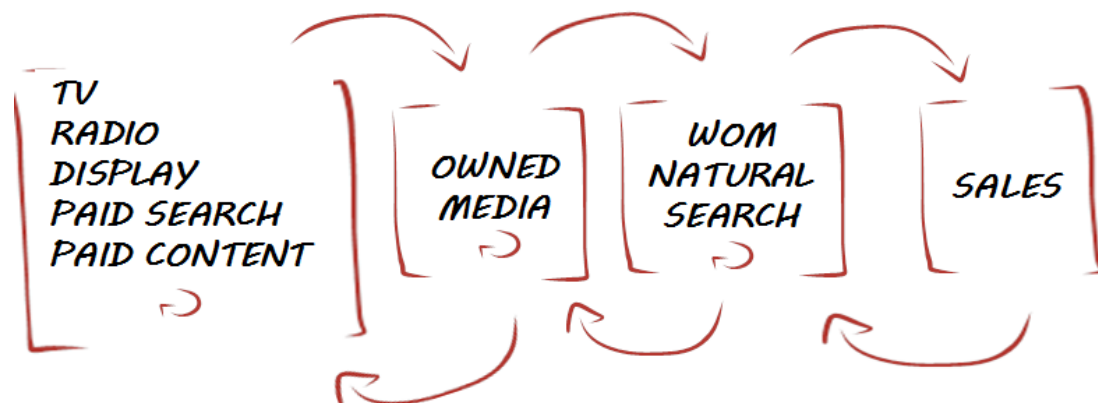
The measurement challenge is not only to evaluate the direct effectiveness of the social element – but also its role in amplifying other media.

Learning from established practices

Social may be new, but the planning process is not and the best way to utilise the new opportunities presented by social is to ground them in what is already known about campaigns and activities. In this context it is useful to think about the planning cycle and the role of short-term and long-term effects.

Stephen King's Planning Cycle outlined below, can be very helpful in asking the key questions that marketers need to think about - not only in terms of their marketing communications activity, but also about the position of the brand as a whole.
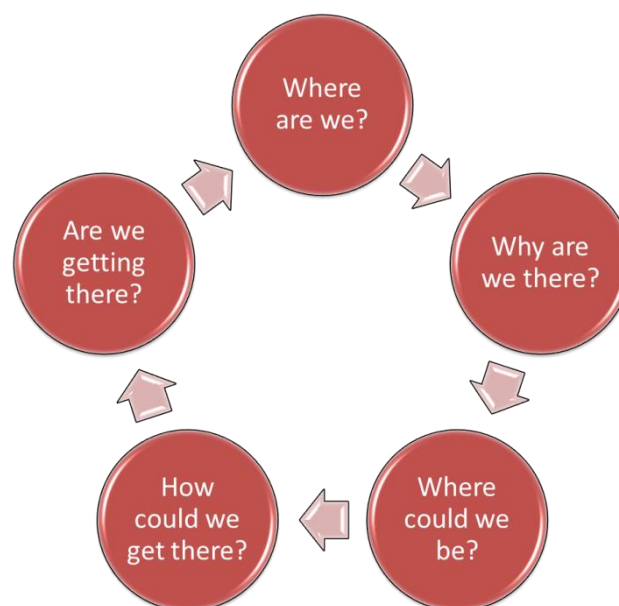
Stephen King's Planning Cycle

In the context of social it is important to be clear about the objectives and the role that social will play. In a perfect world, social will be evaluated with the same tools and approaches as other elements of campaigns, however, there are often pragmatic reasons for recognising differences.

One key issue for the evaluation of marketing is that the long-term effects are not simply the aggregation of short-term effects, and focusing on short-term benefits, such as numbers of clicks, can make it harder to achieve the long-term benefits. Short-term benefits are often specific actions, whereas the long-term effects are more likely to be related to being able to charge more and/or sell more. This presents a real challenge for social where there is a substantial focus on the short-term and an abundance of short-term measurements.

Measuring Social is Hard

This section looks at why measuring social is harder than it initially appears, highlighting the key issues and the ways they are being addressed.

However, it should also be noted that social presents many new and exciting opportunities to measure, for example the ability to review the totality of the campaigns interactions, as opposed to those of a sample, and for that review to be based on accurately collected data, rather than unreliable recall. Similarly, social provides the chance to listen to real, unprompted conversations between consumers, allowing concepts such as sentiment to be measured.

This chapter covers why social is hard and what people are doing to address the challenges, in particular:

- How social is the same, and how it is different
- Short and long-term effects
- Influence
- Homophily
- Causality
- Dynamic frames of reference
- Defining what the activity is seeking to achieve
- Common mistakes

How social is the same, and how it is different?

The obvious measurability of social media initially makes it look quite different, in terms of measurement and evaluation, from other media. However, as well as differences that matter, there are similarities that should be recognised and acknowledged.

The Similarities

Social has similarities to other channels and these similarities should be leveraged when planning the evaluation of a campaign.

- Paid advertising in social media is directly comparable with many other forms of advertising.

- The earned media element of social is sometimes comparable with PR campaigns,

- The direct response of social can be comparable to direct marketing.

- Measuring calls to action can be similar to couponing and some forms of POS.

As with all campaigns, the overall impact on the brand including its perceived value and trustworthiness need to be assessed as well as its effect on the bottom line.

The Differences

The key differences in evaluating social campaigns relate to:

- The measurability of so many aspects of the campaign. Because social is such a measurable phenomenon, problems can arise from having too many choices of data rather than too few, and from having metrics which can be inconsistent over time and across platforms.
- As well as producing measurable metrics from the platforms, social produces vast amounts of naturally occurring posts and comments from consumers and citizens, creating the opportunity to use social media research techniques to explore, ideate, and measure.
- Even more than other media, social tends to be used in combination with other channels, acting directly and indirectly, and being impacted by the other channels. For example, to what extent does TV advertising impact Twitter, and to what extent does Twitter impact TV advertising?

- The need to think about where the social campaign/activity has occurred – is it in paid, owned or earned media? Exposure to earned media and cross-channel activities are often hard to control from a measurement point of view, which can make creating a control sample more difficult.
- The real-time nature of social feedback means that campaigns or activities can be modified whilst still running, which integrates measurement and metrics into the campaign management process. This integration of measurement and management makes issues such as benchmarking and the definition of pre-post much harder.

- The timeline for a campaign can be very different, for example social can last much longer than most other campaigns or activities. Social activity can be 'always on' - especially when used in the context of customer feedback or satisfaction; or as a continuous promotion activity that is not specific to a 'campaign'. This removes the clear 'before', 'during', and 'after' that is a part of traditional measurement methods.

- Although the objectives of a social campaign/activity should always be linked to business objectives, they are not always linked directly to them. Social often used to achieve tactical ends, such as collecting information, encouraging trial, etc. Evaluation needs to be tailored to the specific objectives of the activity.
- The budget for social activities/campaigns is often smaller than for, say, TV and the number of social campaigns/activities may be larger, which means the evaluation/measurement approaches need to be appropriate to the circumstances and budget.
- One of the challenges for social is the lack of knowledge about the issues and the tools available. One study in France found that only 45% of relevant senior executives believed they had a good or very good knowledge of the tools available (Flores, 2013).

Short and long-term effects

The IPA has produced a substantial body of evidence to show that the key impact of advertising is through its long-term effects, in particular when considering outcomes such as profitability. One of the key inputs to the UK's understanding of how marketing works is grounded in the IPA work conducted by Les Binet and Peter Field, and embodied in 2013 in the report The Long and The Short.of It.

However, the immediate nature of social often focuses attention on short-term results and measurements. It is important to recognise that people want short-term measures, and that short-term measures are one of the key attractions of social. At the same time, it is important to ensure the long-term is considered too

The key points about long and short-term effects are:

- Although there are no long-term business effects without short-term effects, the reverse is not true – there can be short-term effects without long-term effects. Long-term effects are not just the accumulation of short-term effects.

- However, there is rarely a long term without short-term results.

- To optimise effectiveness and efficiency, campaigns need a balance of short and long-term objectives.

The key point that Binet and Field make is that relying solely on short-term metrics could be a recipe for long-term disaster. One of the benefits of social is that it provides measurements that allow campaigns and activities to be optimised in real-time. However, the management of campaigns should balance long-term success with short-term success, since they tend to depend on different elements and strengths. Adapting campaigns and activities solely to short-term metrics could sacrifice the activity's long-term prospects.

Evaluating short-term effects

Utilising Binet and Field's works the key messages for evaluation of short-term effects are:

- Activation effects are best measured with immediate behavioural metrics, such as direct response rates, 'click-throughs' or immediate sales uplifts.

- These metrics point in the direction of rational advertising or price promotions, since these tend to produce the biggest short-term responses and volume growth can be quickly achieved.

- However, such strategies tend to under-perform in the longer term. Long-term success is best achieved by emotional brand building effects, which cannot be detected by measuring short-term responses.

- Short-term volume growth can initially be an efficient strategy, as share gain is achieved most efficiently for 1-year campaigns, but not necessarily an effective one i.e. one that drives profitability or other business metrics strongly.

Evaluating long-term effects

Methods for measuring long-term effects are less robust and developed than those for short-term effects, but the key points are:

- The effects of emotional brand building are more subtle, but last longer and grow over months and years.

- These effects cannot be simply measured by short-term response metrics.

- Brand tracking is the established way to measure long-term brand building.

- Emotion-based and fame metrics may also help to rebalance the scorecard in favour of long-term effects.

- Other business metrics such as market share or profit that are usually measured over a period of a year are also associated with long-term effects.

- Share of voice is another, long-term metric, closely related to market share. Comparing share of voice over time against market share is a good predictor of long-term growth, as has been demonstrated in earlier research.

- The most valuable long-term business metric is price elasticity (either in terms of being able to charge more for the same volume, or being able to sell more at the same price). The price premium is important because of its close links to profitability, and it is a metric that only tends to move over the long-term. Unfortunately is price elasticity/premium is a measurement that is complex and often overlooked.

In practice, the brand-building element of a campaign should be evaluated over a period of at least 6 months, whilst activation elements can be evaluated over a shorter term.

Influence

Influence has become a key topic in marketing, and in particular in terms of social. It is also a phenomenon that is being hotly debated. This section starts by presenting the case for influence that has been developed over the last 70 years, and why it can make measuring social hard. The section will then explore why influence is a hotly debated topic and may actually be minimal.

The key ideas underpinning the concept of influence are:

1. Ideas and preferences spread though communities socially, from person to person, as opposed to individuals making independent decisions.

2. Some people are more influential than others inasmuch as others tend to follow their practices and/or advice, these people are called influencers.

3. Marketing to influencers (also known as influencer marketing) is likely to be more effective (deliver better ROI) than marketing to everybody or marketing to, say, demographics.

4. Influence can be measured and this can be used to make influence marketing more effective.

Influence makes the evaluation of social hard for three reasons:

a. The desired effects are intended to occur amongst people who were not directly reached by the campaign/activity.
b. There is considerable dispute about how influence works, and even about the extent to which it exists (with some people suggesting it is minimal).
c. The key issue with influence is incrementality, how many of the people who reacted were 'caused' by the influencer strategy, and how many would have bought/tried/played etc anyway.

Background

Influencers have become a hot topic in marketing over the last 10 to 15 years, particularly since the publishing of *The Tipping Point* in 2000 and *The Influentials* in 2003, although the concept of influencers can be traced much further back, for example to Paul Lazarsfield's analysis of the 1940 US Presidential election. The appendix includes an outline of the history of interest in influencers and highlights key publications and milestones.

The key interest in influence relates to targeting (another topic which is being itself hotly debated). Influence marketing seeks to utilise those who are most influential, and who will carry the message to the wider community. For example, a paid media campaign may be used to reach influencers, with the intention that the influencers create earned media which in turn reaches a wider group.

Celebrity versus everyday influencers

Most of the contemporary interest in influence focuses on the notion that there are influencers in a society, everyday influencers. However, there is another form of influencer, the celebrity influencer, for example people on the Time 100 list or celebrity Tweeters like Kim Kardashian.

Whilst the majority of campaigns do not utilise celebrity influencers, they are very popular for some types of campaigns and activities, and the leading celebrities can command a high fee. For example, The Huffington reported that Kim Kardashian receives US $20K and her sister Khloe $13K per tweet (Kornowski, 2013).

The utilisation of influence for marketing

There are essentially two ways of utilising influence marketing:

1. Targeting marketing at people believed to be influential, with a variety of ways of identifying them.

2. Working with organised groups of people identified as influencers, sometimes referred to as influence panels.

Targeting influentials

Targeting influentials consists of identifying influencers and directing marketing at them. Here the term marketing is used in a very broad sense, for example it might mean trying to get the influencers to be early adopters of a new product, and/or to be knowledge sharers, and/or simply to help promote awareness.

Strategies for identifying influencers include the following:

- If members of a particular platform or group are felt to be influential, then media can be purchased to reach them.

- If people with certain characteristics are felt to be influential, for example people with a certain number of followers, or people with a high influencer score, then usually a targeting system can be found to reach them. The characteristics might be influence scores (such as Klout) but they might also be derived from database scoring.

Influencer groups/panels
Influencer groups and panels, such as members of the Klout Perks programme, are collections of people who are deemed to be influential and who have signed up to take part in advocacy and marketing programmes.

Disputes about the nature of influence
In terms of influence there are two very different schools of thought, those who consider influence to be central to how memes and fashions are transferred and there are those who consider influence to be a mirage or at best a marginal feature. However, most people do not fit neatly into either of the two extremes.

The believers:
If you look at any new trend, fashion, idea you can trace it back and it will have started with a small group of people. In most cases, some of the early adopters will be connected to many more people than others, these tend to be people who appear to be influential inasmuch as the people they are connected to picked up the observed phenomena after they did.

Ed Keller makes the point that there are people who tend to be "*in the center of the conversation.*" It's not that these people are the only ones who pick up on new product introductions, new ideas, etc. But, because they keep up with what's new, like to try new things, have a wider than average social network, and are often sought out for their advice and recommendations, they are people who have greater than average reach or impact through word of mouth.

The final part of the influential proposition is that this pattern repeats itself, i.e. the influencers for one trend or idea are also influential for other fashions and memes.

The disbelievers:
The disbelievers have shown that many of the patterns observed in the data can be replicated without any assumptions of influence. This replication is normally achieved by either modelling actual data or creating artificial networks. These models tend to assume ideas are spread socially, but do not assume that some people have a special ability to 'influence' people.

This model assumes that we are all influenced by other people, but disputes that there is a class of person who can usefully be thought of as 'influencers'. For example, in a shoal of fish or flock of birds, each animal is sensitive to its immediate neighbour and will copy their movements, allowing the shoal or flock to move in unison. However, the 'influence' is solely due to position. At a different time, the fish/birds are in a different pattern, so the copying changes.

As a consequence, the disbelievers do not believe that it is better to target influencers than, say, a random set of people (or not substantially better, depending on the degree of disbelief).

The limited influence model
The leading voice at the moment in the area of measuring influence and homophily is Sinan Aral, for example his paper *Identifying Influential and Susceptible Members of Social Networks* published in Science, June 2012. The key conclusions of Aral's work are:

a. Influence exists, is measurable, but is smaller than is generally believed.
b. That models that do not take homophily into account tend to overestimate influence.

---

c. That influencers and influence patterns for one product category may not hold for other categories.
d. Seeding, giving products to 'influencers', should be restricted to a very small proportion of the universe (in one study the maximum suggested was just 0.2% of the population).
e. That providing incentives to influencers tends to have a larger effect than seeding, and that combining them works best, but is still of limited effect.
f. To assess the impact of using influencers it is necessary to assess what would have happened anyway. Many people would have adopted the behaviour under review without any influence being offered or assumed.

Pragmatic Approaches

Professor Alex Bentley suggests a pragmatic approach:

"*The definition of an 'influential' is often tautological - a person who picks up a new idea early and spreads it to many other people. The empirical question in all cases is whether these people fit any fixed categorical description, and whether influence ever travels the same path twice. For me it depends on:*

1. *The longevity or tradition of the behaviour and the propensity for recognised experts to exist. Is this a Nobel Laureate in Economics giving a lecture or someone using a Samsung phone on the street?*

2. *The degree of 'information overload' that may overwhelm people's ability to focus their social learning toward prestigious or expert individuals.*

3. *The scale of the phenomenon - is this a small community or a population?*

One example might be to consider the role of 'influence' in the models of traffic, flocking and pedestrian dynamics, which focus not on trying to trace the path of individual influence from one person-particle to the next, but on the architecture of the space, on adding a column in front of a doorway, the effect of merging, fluid flow, etc. This is the appropriate scale. Now, that said, you might have a situation where you can identify someone, like the person in your building during a fire drill, who wears the yellow vest and directs people out.

So, I think the case-specific question is whether influentials exist in a given situation. If so, give them a yellow vest. If not, build a column in front of the doorway, so to speak."

Best practice advice when using influence marketing

1. Decide on whether the campaign will target influencers or will work with one of the influencer groups to co-market/co-promote the campaign.

2. Determine what you believe would happen if you selected a random group to market to, and benchmark against that.

3. In terms of benchmarks, consider what the extra costs are (if any) of working with influencers.

4. How you will determine success, in terms of gains over and above the non-targeted benchmark? How will the ROI be calculated?

Homophily

Homophily is a counter proposition to that of influence, and indeed an alternative to much of the thinking behind targeting and causation. Social scientists have noted that similar people tend to cluster together, as in *Birds of a feather flock together*. What may look like influence can in fact be lots of people responding in the same way to some external stimulus. This phenomenon is known as homophily.

As social media has grown, people have become more interested in how ideas propagate, through peer-to-peer networks and/or through mass media. The temptation is to see viral communication everywhere. However, clusters of similar people tend to congregate together and may look like they are influencing each other, but are in fact responding to the same external stimulus - seeing a 'Quit Smoking' ad on TV, for example. In some academic studies, over half of what looks like a cascade is actually 'homophily' (Aral et al 2009, Aral 2011). The implication is that TV and other mass media may be more important than at first sight (Watts, 2007).

The difference between influence and homophily is of particular importance when determining how to target marketing. For example we might notice that if one person in a social network buys something, other members of that network also buy it, then we might draw one of two inferences and design two very different marketing campaigns.

1. Some people are influencing others, so we should find and target potential influencers.

2. We might assume that people with similar tastes tend to form networks, in which case we should target people in the same network as somebody who buys something, or who is similar to people who buy it.

This is the difference, from a targeting point of view of influence and homophily. At the moment the leading expert in the field of untangling these two positions is Sinan Aral, who has shown that influence is measurable, but smaller than is generally assumed, and that failing to account for homophily leads to influence being greatly over-estimated.

Causality

"*Correlation does not imply causation*" is one of the most frequently repeated statistical tropes, but whilst correlation is broadly understood, the nature, complexity, and importance of causality is less widely discussed and appreciated.

Why Causality Matters?

Understanding 'why' something works is the key to repeatability, learning, and scaling. Most importantly, causality is a cornerstone of prediction (Gelman, 2010). This is just as true in the age of big data as it has always been, "*a theory-free analysis of mere correlations is inevitably fragile*" (Harford, 2014).

The key to replication is being able to tell stories, to describe how things work. The ability to link ideas causally is likely to be the dominant approach in marketing for many years. Machine learning approaches have been available for some time, for example neural nets, but they are not widely used because they don't generate clear causal stories. Without causal stories it is much harder to create new material, new campaigns, and new activities.

In a world with ever more granular decisions, causality is even more important to understand, because of the need to provide advice on how to respond quickly. Historically, campaign evaluation was a 'justifying the budget' conducted after the campaign ended. Now, more decisions are 'live', and granular, facilitating intervention and real time management. Understanding causality is key to speed, particularly where there are complex explanations

Causality and ROI can be very hard to prove conclusively, especially with the limited resources available to process social data and information (Lewis & Rao, 2013). The two key issues are often:

    a. The need to make explicit assumptions.
    b. Adopting a disciplined approach (rather than adopting a fishing approach).

One key approach is to consider consumer decision journeys (Court et al, 2009). The decision journey recognises that the traditional funnel paradigm is too narrow and linear for the multiple touch points that have accompanied the explosion in product choice and digital channels. Because

the decision journey is not a linear model the measurement of causality becomes much more complex. Feedback loops and compound effects need to be accommodated.

There is a school of thought that in the age of Big Data causality does not matter (Anderson, 2008). However this view was demolished by Nate Silver (2012) – and most commentators believe it to be nonsense.

Why is causality hard to measure in social?

Digital and social media might be the most measurable media, but it can also be the easiest to misinterpret. While its interactivity – constantly asking people to make decisions and choices - creates many useful metrics. Interactivity can be the reason that it is difficult to understand what influences final decisions. Unpicking cause and effect in digital media is riddled with nuance.

One example of where measurability can struggle to identify causality is last click attribution, i.e. the faulty assumption that the credit for an action can be ascribed to click that leads to an action. For example, Finger and Dutta (2014) quote the CEO of a media company who complained "*When we publish an article on how good olive oil is for your overall health, our clients will go online and look for olive oils. They will most likely buy something we have recommended in our publication. However, the search engine will get the money, not us*." Finger and Dutta assert that in cases like these "*search engine advertisement has very little to no causal effect on the intent to buy*."

Establishing and measuring causality requires careful construction of counter-factual scenarios, i.e. what would have happened without the marketing intervention. The gold standard in causality is randomised controlled trials. This approach is good because it is clear that the only difference between the outcomes of different trials is the impact of the stimulus, however, clean experiments in social are hard to construct, because it can be hard to create cells who are not exposed to specific messages in the case of earned media and often when multiple platforms are used.

Three main issues with causality in measuring social are:

1   Selection and targeting bias

Even with careful design, exposure to media and marketing is not random, especially when taking multiple channels and earned media into account (although paid media can be more controllable in social than in other media).

However, the deeper problem is that behaviour online and exposure to marketing is often commonly driven by intent. For example, the more somebody wants to buy a product the more likely they are to see related media about it. For example, if somebody intend to buy a Canon camera from Amazon, they are more likely to use Canon related search words, to see Canon display ads (particularly with re-targeting) and even to be a follower/fan of Canon Facebook. But none of this has influenced their decision to buy.

This 'selection bias' matters a lot. In a recent study, it was found that naïve measurement can be biased to the tune of about 1000%! (Farahat & Bailey, 2011). Variables such as clicks, reads, plays, and posts, should be treated as both inputs and outcomes.

The people who respond 'virally' to campaigns are even more likely to be a biased sample. Best practice should include test and control measures wherever possible, even with econometric models or with matching techniques.

2   Synergy and interactions with other media

A/B tests try to isolate the impact of individual marketing actions. But in reality marketing often works synergistically. Marketers are interested in how social media and search or social media and TV work together rather than in isolation. For example, a Twitter interaction with a TV programme can provide re-targeting opportunities. It is important to make sure that the 'credit' is distributed to the right media channels and content so that budgets can follow accordingly.

More generally, the relationship between media is seldom linear. Not only are there diminishing returns to scaling up investment, but there can be critical mass impacts (e.g. people respond better to social media campaigns when the brand achieves a 'threshold' level of salience and brand equity). Word of mouth marketing relationships are typically non-linear, include multiple variables (such as quality, exposure, and length) and involve feedback, e.g. liking something encourages buying, buying encourages liking.

3   Homophily
Homophily suggests that similar people do similar things and that this can look like cause and effect. In essence this comes back to the counter-factual issue, if the campaign or activity had not been run, how many people would have bought it anyway.

Unpicking causality in social media - practical steps
Whilst there are few situations where there is a perfect solution to establishing and measuring causality, there are a number of practical steps that can be followed.

- **Benchmarks and targets.** The most common mistake in digital measurement is to simply quote the number of followers/fans, views, clicks or impressions. The metrics produced need to be compared with something to make them more tangible. Setting targets before a campaign is a good idea. Starting with a prediction is the first step to thinking about causality. Comparing the scores with benchmarks is also a good idea, using earlier campaigns or other current campaigns as the benchmarks. In many cases the platforms may be able to share norms.
- **A/B tests**. In A/B testing two cells are matched in every way bar one, and that one is the item being tested. For example using two executions, or two targeting options, on a specific platform, such as Facebook, Twitter, or LinkedIn. If the two cells are matched, and if there is no contagion between the two cells, then it is reasonable to assume that any major differences are causally linked to the stimulus.
- **Stating all relevant assumptions**. In order to measure causality it is usually necessary to define the variables and the relationships that are being assumed. For example, what are the time periods, whether earned is being included, and what variables are being treated as endogenous.
- **Modelling direct and indirect effects**. Causality is rarely binary or linear. It is important to look for direct impacts, thresholds, and importantly interactions. Often, the best way to capture complex dynamics is through econometric modelling, a topic which is covered in the appendix.

Dynamic Frames of Reference

One of the challenges in measuring phenomena in social media is created by the fact that almost every feature of social is changing. Measurements tend to require consistency, to allow models to be created from observations and allowing predictions to be made and tested.

In the context of social, the number of users is changing, the competitive set is changing, the usage patterns are changing, and the ways that companies are utilising social is changing. One of the outcomes of this process is that most measurements in social do not have an absolute value, requiring comparative measurements to be utilised.

One of the key requirements created by the dynamic frames of reference is the need to create and update benchmarks. Benchmarks can be multiple activities for a specific brand or service or they can be measurements taken for several brands and services.

When a medium is changing quickly it is harder to create benchmarks, for example, as the number of people using Twitter grows and the number of tweets grows it is harder to benchmark one campaign

against another. Similarly, click through rates dropped from the 1990s onwards, making comparisons difficult.

Google Trends benchmarking

Google Trends reports the incidence of people searching for term on Google over time. However, since the number of people using Google is growing a method had to be found to standardise it. Google does this by allowing the user to specify the terms to be searched, the area (e.g. UK) and, and a period of time. Google then calculates for each period of time, for example a week, the total number of searches and for the items searched their percentage of the total. Once this has been done for all of the time periods the largest percentage, for any one term in any one time slice, is set to 100 and all the other points scaled to fit.

However, this system has two major drawbacks:

1. The numbers only relate to that collection of terms, for that time period, for that region – there is no generality to the numbers.

2. If something else is happening in a particular month, the numbers will appear to fall, even if the absolute numbers have not fallen. For example, many terms appear to dip in the run up to Christmas, which may not reflect a dip in the number of searches for that word, but could reflect an increase in searches for Christmas related items.

Google Flu Trends

One of the most famous uses of social media monitoring, since 2008, has been Google Flu Trends. Google Flu Trends measures the incidence of people searching for words like cough and cold. This process has been generalised by Google with tools like Google Correlate and Google Insights.

However, a recent article in New Scientist showed that despite Google's early success in predicting flu, its accuracy has declined over time (Hodson, 2014). The article makes the point that the changing nature and variability of social media has reduced the ability of Google to predict the likely spread of flu in the USA.

Defining what the activity is seeking to achieve

The traditional way of measuring campaigns has been to focus on ROI, and ROI remains important in the context of social. However, there are other endpoints that a social campaign might be seeking to address and these need to be identified too.

Non-ROI goals

A social activity might have a goal such as increasing the likelihood of applying for a job. In a perfect world this goal will have some link with the final objectives of the business, but the link between the specific social activity and the final ROI may well be too long to be measured in simple ROI terms. In these cases intermediate goals are used. For example:

- Assumption: attracting the right prospective employees is good for the long-term bottom line of the company in the long run.
- Set intermediate goals that measure changes in the number of good applicants.
- Evaluate social campaigns against the intermediate goals.

In the example about the metrics would probably need to include the number of suitable applicants and some measure of how easy they were to handle. For example, if the activity led to a 100% increase in applicants and a 10% increase in good applicants, that might not be effective. Similarly, if a new campaign led to a large number of interactions being required with potential applicants, that might not be effective either.

Efficiency versus Effectiveness

One of the key issues surrounding ROI is the difference between efficiency and effectiveness. Overemphasising either efficiency or effectiveness can generate sub-optimal results.

Efficiency focuses on the ratio of the return to the investment, which is one of the standard definitions of ROI. The key problem with overemphasising efficiency is that it tend s to favour options with lower levels of spend, which can result in the most effective campaigns not being run (because they may be less efficient).

The effectiveness of a campaign refers to the extent to which the objectives are met or maximised. In terms of ROI, effectiveness tends to refer to the total return, as opposed to the ratio of the return to the investment.

ROI - what can be measured?

The main challenge in calculating ROI is the 'R', i.e. the return. To calculate the ROI a Return has to be defined and it has to be possible to link it to the investment. A social media campaign might intend to improve engagement with a Facebook page in order to get other people to engage with the brand which will make people more aware of the brand and better disposed towards it, leading to more sales. The increase in sales is a 'Return', but there are often too many external factors to clearly link the return to just the social campaign. In terms of social, one of the factors that makes the calculation of ROI more complex is that social is usually employed as part of an integrated campaign/activity, so typically we want to understand the ROI of something like Twitter plus TV, not just Twitter in isolation.

Several brands have found success by creating intermediate goals. For example if a brand can show that increase in customer satisfaction or positive sentiment is associated with an increase in sales, then these measures can be taken as the 'Return' that the social activity is measured against.

As well as increased revenue, 'Return' can be accrued by reducing costs. For example, if customer services, or branding, or the collection of insight can be managed at a lower cost through the use of social, there will be a return, against which the costs can be assessed.

The I in ROI, the investment, can also present some difficulties in the context of social. Thinking about the POEM model, the cost of paid media is relatively straightforward, but the cost of owned and earned can easily be underestimated, which in turn can result in the ROI being over-estimated,

The O2 case study uses both intermediate goals (they have calculated the value of one percentage point of customer satisfaction) and cost reduction (using social to reduce the costs of customer support). It is likely that the future of ROI calculations will feature intermediate goals and cost reduction as key elements.

Common mistakes

There are a number of errors to be alert to when evaluating social, including:

Perverse incentives. A perverse incentive is created when pursuing a target results in an undesirable outcome. For example, pursuing more visitors can result in more inappropriate visitors and potentially fewer relevant ones, even in absolute terms.

When dealing with social many potentially interesting metrics can generate perverse incentives. For example, if a brand creates a truly engaging campaign or activity it will often acquire more followers/fans – and the growth in followers/fans is an indication of its popularity. However, if followers/fans are set as the objective, e.g. as a KPI, then brands and organisations may start to look for easy ways to boost the number of followers/fans without any regard to whether or not they help the brand at a deeper level.

As an example of a perverse incentive, Finger & Dutta (2014) cite a client who made CPO (cost per order) their key metric. Initially, this seemed to work, the CPO which was already low fell even further. However, they found their sales were falling too. The campaign was being targeted mostly at existing customers, who were likely to click, which helped achieve lower CPOs. The unintended consequence of prioritising CPO was to reduce the size of the pool the campaign targeted to an adverse level.

Metrics that are easily gamed. In this context gamed means that a high score can be achieved without the underlying phenomenon of interest increasing? For example, if followers/fans/plays can be acquired easily/cheaply from a third-party supplier, the metric is easily gamed. There is clearly a link between perverse incentives and metrics that are gamed. The difference relates to focus, with perverse incentives the focus is on how it will impact the way the business is run, with easily gamed metrics the focus is on the integrity of the measurement.

Vanity metrics. Vanity metrics is a term that is used for metrics that can make a brand or campaign look good, but which may have little impact on the performance of the brand (Ries, 2009). The number of visitors, followers/fans, or downloads can be considered vanity metrics if they are not targeted at a specific purpose and linked to key outcomes. In general it is not the metric itself which makes it a vanity metric, it the way a metric is selected and used that creates a vanity metric.

One problem with measures like plays, followers/fans, or likes is that they may not reflect current levels of interest or engagement. Few people unfollow, unlike, or remove comments, which means that the historic number of connections does not necessarily relate to what is happening at the moment.

Fake followers/fans. There is growing business in creating false followers/fans, typically utilising low-cost economies and/or bots. These make measurement problematic. Measures such as influence and advocacy can be thrown badly off track if fake followers/fans are included. In November 2013 StatusPeople.com published an investigation of Katy Perry, who has nearly 50 million followers, and estimated (they use a sampling method) that a majority of her followers were either inactive or fake. However, the picture for one celebrity does not necessarily give a representative picture and other studies have produced lower estimates, with Business Insider quoting figures of about 5% from both Facebook and Twitter (D'Onfro 2013).

The providers of metrics and in particular the platforms themselves are aware of the problem and are constantly updating their systems to try to eradicate the bogus elements.

The wrong number of KPIs. If too many KPIs are selected the real meaning can become obscured. A large number of KPIs will usually mean that there is too much correlation, which can lead to double-counting. Also, with too many KPIs something in the mix is likely to look good, leading to people highlighting it. Having too many KPIs is more like trawling than measuring - and suggests the organisation does not understand the real drivers of the business.

Too few KPIs is likely to result in something being missed or produce a finding which is so narrow that it struggles to explain the wider picture.

Benchmarking problems

Few activities in social media have an absolute value. It is hard to say what 1000 views, 10,000 Tweets, or 20,000 shares means, unless the numbers are benchmarked against other similar brands, campaigns, and activities. The benchmarking process should be started at the campaign planning stage, by identifying relevant comparators.

Failure to set a target for the key metrics is a common failing.

How to Measure Social

This section consolidates the advice and input received from industry leaders in terms of what is currently seen as the best way to measure social campaigns and activities. We have not called it 'Best Practice' since the field is too new and dynamic to attach that sort of superlative to the advice. However, it does appear to be the best advice on 'how to measure social' that is currently available.

This section covers the following:

- Designing an Evaluation Strategy and Process
- Five point plan for 'baking' evaluation into social
- Data, metrics, and KPIs
- The platforms
- Selecting the metrics
- Evaluation methods and approaches
- Using market research to evaluate social

Social media research, for example the monitoring of discussions in social media and the creation of sentiment scores, is also used to assess campaigns and activities, and this is the subject of the next chapter.

Designing an Evaluation Strategy and Process

Measurement does not just happen; it requires a framework, it requires planning, and it requires the selection of the right metrics. The measurement of traditional media has been developed over decades and has been subject to a high level of testing. By contrast, social is relatively new and highly dynamic, which means that there are few established rules and practices.

Because social is less well developed, there is a greater need to design a measurement strategy, rather than adopting an existing approach or plan. Most social campaigns and activities operate in conjunction with other media and this factor needs to be an integral element of any evaluation.

Establishing expectations and a framework

As with the measurement of any other form of media, measurement in the context of social media requires a framework and it requires an understanding of what can and can't be done. Although some elements of social media are different, the fundamental rules still apply. The practices for social media should build on what has already been learned about how media works.

The expectations and framework should define the goals of the campaign/activity and check they are aligned to the business strategy/objectives. Because social is seen as undeveloped and evolving campaigns are often undertaken because there is a perceived need to 'be doing something'. However, even in these cases targets and predictions should be created, to allow the campaign or activity to be adequately assessed.

Benchmarking and targets are key to the evaluation of social media campaigns and activities. Few social media metrics have any independent, absolute meaning. Fifty thousand 'likes' or tweets is hard to interpret, but when compared to two thousand or two million 'likes' or tweets there is, at least, some context. As the IAB have said, "*Without comparative benchmarking most metrics are completely meaningless*".

The use of KPI's requires a target, which should be set in advance. Benchmarks are common method of setting targets, but other methods can also be used, for example forecasts based on market mix modelling. The key feature is to set realistic expectations and performance measures.

Planning ahead

As with any measurement of campaigns or marketing activities, the campaign objectives and the items to be measured must be set before the campaign starts. Specifically, the metrics need to be identified and the methods of collecting the data should be agreed.

With social media there can be a temptation, since the data for earlier time periods often exist, to assume it is OK to wait until the campaign is underway or even finished before looking backwards to find evidence of its impact. This is flawed and will often lead to the wrong interpretation. The presence of external/unrelated factors mean there will usually be something in post-reviewed data that looks positive, and trawling for information will generate spurious findings and correlations.

Try to plan the media to facilitate measurement, for example by not scheduling everything to happen at the same time. The ideal is to create experimental cells, where different people will be exposed to different executions and combinations.

Being SMART

The evaluation process should embody a SMART approach, i.e. Specific, Measurable, Achievable, Relevant, and Time bound. The evaluation should not confuse objectives with metrics or KPIs. The objectives are what the campaign or activity is aiming to achieve. The metrics/KPIs are measurements that should indicate the extent to which the objectives are being been met. It is important to recognise that there is no single best metric and that the best metrics for the last project may not be the best for the next project.

Select measurements that are quantifiable, facilitate action, and linked to business outcomes. One key element in making the metrics useful (for both action and evaluation) is to limit the number of KPIs, ensuring they are easy to understand and efficient to collect.

Set aside time, budget (perhaps 1% to 2% of the campaign spend, or 5% for smaller projects), and resources for analysis. There is a perception that the data from social is free. Whilst the acquisition of data is sometimes free (and sometimes not), the proper analysis of the data takes time and money. Investigate what analysis options are available from the platforms and providers.

Look beyond just pre-post. Look for relationships over time, using data from before, during, and after the campaign/activity. Think about other factors, such as: seasonality, homophily, other brands, other campaigns, ongoing activities.

Distinguishing between efficiency and effectiveness

Efficiency is a ratio bases approached, typically with the desired outcome (e.g. sales, leads, registrations) as the key number, and the 'cost' (for example money, resources, or time) as the base. Efficiency is, effectively the same as one of the measures of ROI. If you obtain $4million of value for a spend of $1million, then the efficiency/TOI is 4 million over 1 million, or 4.

Effectiveness is a measure of how much the activity delivered. In the example above the effectiveness would be $4million (or arguably $4 million minus $1 million, i.e. $3 million).

A good campaign is both efficient and effective.

Integration with other elements of the campaign

Whilst some social campaigns or activities are 'pure play' (i.e. they are only social) many are part of a broader picture, utilising a number of channels. Where social is being integrated into a broader campaign the evaluation needs to operate at three levels:

1) The impact of the social media.
2) The total impact,
3) The contribution of social media to the total picture.

The third element in this process usually employs an attribution model or market mix modelling - which is covered more fully in the appendix.

The metrics for social have two roles, the first is to help optimise the campaign, and the second is focused on evaluation and measuring. Some metrics will be used for both roles, others will be chosen for just one role.

Continuous measurement versus ad hoc campaigns

As with other marketing activities it is possible to conduct continuous measurement and/or measurement targeted at specific social campaigns/activities. The key issues for these two types of measurement are:

- **Continuous** metrics should seek to measure expected changes and at the same time have the capacity to pick up unexpected changes and effects. One examples of a metric that can help identify changes that were not expected is sentiment.

- **Ad hoc** campaigns typically require that the key changes that the campaign is required to create be identified in advance. Once the key changes are identified, baselines can be created. In some cases the objectives might need to be tweaked to ensure that the results are measurable.

One similarity between continuous and ad hoc campaigns is that the real-time and granular nature of social media measurement means that the management of campaigns can often be tweaked or amended in the light of the social media feedback.

Advertising versus services (outgoing versus incoming messages)

The measurement of campaigns, for example advertising, will focus on the goals the campaign was designed to impact, plus the measurement of core brand/organisational attributes. By contrast the measurement of incoming messages relating to service provision and the general social media buzz about the brand needs to be more open-ended.

The measurement of services and general buzz will often include the number of people talking about a brand in social media, sentiment analysis, reviews, and referrals. The BT case study quoted earlier provides a good example of how service delivery and social can be integrated, how they can be part of a wider campaign, and how they can be measured.

It is important to note that customer satisfaction measurement by social is a developing field and needs careful analysis. Twitter, for example, may achieve low scores for initial satisfaction, because it can be used as a "complaint" route. Customers who are unhappy and who Tweet can, of course, be converted to brand champions if handled in the right way, but the conversion may not be handled or expressed through the same channel and therefore the channel's effect may be misattributed.

Five point plan for 'baking' measurement into social

To properly evaluate social media marketing and activities, the measurement process should be designed in from the start. Waiting until the end of a campaign to determine how to measure the campaign can fail either because the right data have not been collected or because there are simply too many possible interpretations of what happened.

The IPA has generated a five point plan to help ensure that measurement is baked into social campaigns and activities. Note, these are not five linear steps. Organisations and their uses of social, are very varied, so different projects might enter the process at different points and may travel through the list more than once to ensure that the opportunity to evaluate the campaign is maximised.

1. What is the campaign/activity designed to do? How does this link the wider business or organisational needs? Not all social campaigns are intended to link to broader business objectives, which is fine as long as the aims are defined. The measurement of a campaign should be linked to the aims of that campaign.
2. Why social? Why has social been chosen? Is it being used on its own or in addition to other components? Which channels are going to be used? Including a consideration of the role of Paid, Owned, and Earned Media.
3. What decisions will be made on the strength of the evaluation? Are the metrics to be used at the end of the campaign to assess it, or are some to be used dynamically during the campaign/activity to manage the implementation? How would these decisions be made otherwise?
4. What are the most appropriate datasets and metrics? How will they be collected? The key considerations are:
   a. Audit metrics, e.g. what was sent out and who was reached, including virality.
   b. Resonance metrics, e.g. engagement, attitudinal, customer satisfaction, reputation effects
   c. Response metrics, e.g. sales, customer acquisition and other calls to action.
   d. Metrics that differentiate between the channels and between social and non-social components e.g. TV and Twitter, Search and Social.
5. Designing the evaluation process. Determining how the data, metrics, and objectives are going to be analysed to understand the effectiveness of the campaign or activity, and to assess incremental value. Including assessing where the value of social comes from.

Steps 1 and 2 help determine the nature of the campaign and the balance between the channels, for example how much of it will be social, and within that what specific channels are going to be used. Steps 1 and 2 should also generate the hypotheses that the evaluation will seek to test and evaluate.

Step 3 determines the priorities and timeliness of the reporting. Steps 4 and 5 determine what will be collected and how it will be analysed.

Data, Metrics, and KPIs

With the growth in platforms and the proliferation of metrics the challenge is to select the right metrics for a specific purpose. No single book or document can list and review all the metrics available, although Stephen Rapport's *The Digital Metrics Field Guide* does a good job of reviewing 197 key metrics on 12 key fields.

The key issues surrounding the choice of data, metrics and KPIs are set out in the sections below.

Source of metrics

There are four key sources of metrics, each with their own merits and issues:

- The platforms, such as Facebook, YouTube, and Twitter.
- The client, when owned media is being used.
- Third parties, such as ComScore, Klout, and Google – along with market research companies and social monitoring companies.
- Participants, for example location data gathered from mobile device users.

A taxonomy of metrics

The metrics used with social can be divided into the following broad categories:

- **Followers/Fans:** including fans, contacts, followers, for example people who have 'followed' a Twitter profile or 'liked' a Facebook page. As a metric these tend to be easy to count and have initial face validity. However, they are sometimes gamed, and tend to reflect the sum of historic interactions as opposed to current activities.

- **Volume:** a measure of the amount posts, including tweets, shares, and re-tweets. Volume is relatively easy to measure, but may not reflect the effectiveness of activity. The key weakness is that the amount of material produced is not necessarily directly related to its impact, and even the total amount of buzz created is not the same as the quality or effectiveness of the messaging.

- **Reach:** is a measure how many people were able to see or interact with the content produced by a campaign. Its strength is that it produces a measure of breadth to complement a measure of volume. A key weaknesses is that opportunity to see does not necessarily imply impact.

- **Engagement:** is usually a derived measure, based on who is sharing, linking, commenting on, re-tweeting, liking, playing, viewing, or favouriting. The concept of engagement is widely accepted as a key measure of social. The key challenge is defining which elements should be combined to form a specific engagement metric. Another challenge is creating measurements that are easy to understand and consistant across platforms and over time. Most of the platforms have a preferred measure, as do many third parties, for example Klout.

- **Time spent:** (e.g. playing or viewing) is a useful measure for content that requires longer to consume, such as videos, games, or longer posts. The key strengths of time spent include that it can often relate to one of the objectives for a campaign and/or be a component of engagement. There are relatively few weaknesses with this measure, however, if the material is supposed to be intuitive and helpful (such as an interactive map) the intention might be that people should be able to use it quickly, and potentially that they would not need to come back to it, for example the material was a 'how to' video.

- **Referrals:** some tools, for example Google Analytics, suggest where traffic is coming from. The key strength of referrals is that it indicates which links or actions are associated with people interacting with the target material. The key weakness is that it can encourage analysts/users to attribute the final outcome to the last step in what might have been a long chain of causes and effects.

- **Sentiment analysis**: tends to measure whether comments are positive or negative. However, sentiment analysis can be used more subtly, for example evaluating whether campaign is producing responses that are 'on message'. The key advantage of sentiment analysis is that it goes beyond how many people are talking about something towards what are they saying and what they mean. The main negative is the wide degree of scepticism about automated sentiment analysis – although this is being addressed in a number of ways, explored later in this chapter.

- **Share of voice**: reflects the relevant mentions divided by all mentions in the category. The term has a least two distinct uses. When talking about ad revenue models, share of voice is the ratio of the brands relevant impressions to the total number of relevant impressions. Amongst social media researchers it tends to be the ratio of the relevant brand mentions as a proportion of all the relevant brand mentions. The key strength of share of voice is that it is not based on absolute values. Share of voice benchmarks the activity against a wider category. The key weakness is similar to its strength, a good score can be a reflection that there was little to compete with at that time, and a less good score could reflect that there was something else being talked about, such as a major news story or event seasonal issues,

such as Christmas/New Year. Share of voice is best used in combination with measures such as reach or volume.

Validity and reliability

The chosen KPIs need to have the appropriate reliability and validity, covering the following key points:

- Validity: does the measurement correctly reflect what has happened, including internal and external validity.

- Internal validity: is the change being measured caused by the underlying phenomenon we are seeking to assess, for example, is the change in 'liking' caused by the campaign?

- External validity: can the finding be generalised to other situations, for example if the campaign is rolled out on a larger scale would the same effects be observed?

- Reliability: if the same measurement is run twice, measuring the same phenomena, will it provide the same results?

- Accuracy: can the data be collected with relative accuracy?

- Cost: can the data be collected at a reasonable cost?

A measure can be reliable without being valid, if it is measuring the wrong things, or with a systematic bias. And it can be valid without being reliable if it is unstable or prone to measurement error.

The Platforms

There are a wide range of platforms, for example Facebook, Twitter, LinkedIn, Pinterest, and YouTube. Each of the platforms has its own strengths in terms of the metrics it can offer. The metrics from the platforms can be enhanced in two ways:

a. Working with the platforms to track behaviour, for example by using cookies. The range of tracking options changes over time and the different platforms have different options, some going well beyond cookies and email addresses.
b. Third parties, such as Datalogix, comScore, Nielsen and others, offer addition options, in terms of metrics, identification, and tracking.

One of the key things to keep in mind in terms of the platforms and the measurement of social is the potential to gain detailed inputs from paid and owned material – from the POEM model. When working with a single platform a variety of options are available to target who sees what and to track exposure and in many cases behaviour. However, earned media is much less amenable to control and targeting.

Selecting the metrics

Selecting the metrics for a specific campaign or activity is likely to be a bespoke process, there is currently no simple template that can be applied. The five step plan outlined earlier in this chapter will help ensure the key items are covered.

Step	Role	Elements
1	What is campaign designed to do?	List the macro and micro objectives. Macro objectives tend to be business objective such as sales. Micro objectives relate to how the campaign is intended to work, such as downloads, registrations, or sharing. Each needs to be covered by relevant metrics.

2	Why social? What channels?	List the channels to be used. Paid and owned media channels tend to be self-evident, but earned media can be quite wide.
3	What decisions?	The decisions that are made drives the selection of metrics, and the timing of the decisions is a major factor.
4	Select the metrics	Create a matrix of what is possible with the objectives and decisions, to determine an optimal set of metrics. Small enough to be manageable, large enough to cover the needs.
5	Design the evaluation and reporting	Design the evaluation, including the method, budget, timing, and reporting. With social an integrated dashboard showing real-time feedback is often expected.

Evaluation Methods and Approaches

The evaluation of social can be broadly divided into three categories:

- Managing a campaign or activity.
- Formal evaluation of causes and effects.
- Pragmatic measurement.

Evaluating the impact of social media is not easy. With traditional paid media, there are clear campaigns with control over who was exposed, how and when, ideal for measurement, if not always marketing effectiveness. This is less the case with social media, especially earned media. Consumers responding to social campaigns and activities are more self-selecting, are likely to have a predisposition to the brand and their engagement is often driven by external factors (including paid advertising). Within a single channel there can be a high degree of targeting and control in terms of owned and paid media, the complexity arises between channels and in particular to earned media.

When evaluating social media we are often interested in how it works in combination (e.g. Twitter and TV or Search and Facebook), rather than how it works in isolation from other factors. Consequently, the choice of measurement technique is very important. As with any aspect of measuring communication effectiveness, it is important to remember that while some approaches are certainly better than others, the best advice is to look at a variety of methods in the context of clear hypotheses of how we think the communication is working.

Formal evaluation of campaigns

There are a variety of measurement approaches that are applied to evaluating the effectiveness of social campaigns and activities, the key ones are:

a. Market mix modelling (often referred to as MMM) – is considered the 'gold standard' approach to measuring communication.
b. Experiments and A/B testing, for example fusing third-party data with exposure to different elements.
c. Event or individual/household level modelling, which tends to fuse elements of the previous two methods, to provide alternative approaches to the data and campaigns.

The formal evaluation of campaigns is covered in the appendix, address issues such as what sort of data are required and the situations when one method might be preferred over another.

Pragmatic measurement

In many cases social campaigns/activities may be too numerous or have too small a budget to warrant formal evaluation. In these cases pragmatic approaches are adopted. In particular:

1. Identifying chains or connections. For example, if evaluating a campaign to raise the intention to visit a destination, we might measure: the number of posts, the number of people engaging with the posts, and the number visiting tourist websites. In a formal approach there would typically be an attempt to check that the people visiting the websites had experienced the campaign, but in a pragmatic approach the check is often at the aggregate level.
2. Focusing on outcomes. For example, a social campaign may be run to promote a snack. If the sales go up the campaign is assumed to be successful, if the sales do not go up it is assumed not to have been successful.

Pragmatic approaches adopt an informal Bayesian approach, if I do A and X happens, and I do C and nothing happens, then, all other things being equal, there is a better chance that A is beneficial than B. However, this thinking can be flawed. The main two reasons it may be flawed are: a) It does not consider the counterfactual, what would have happened anyway, i.e. if A and B had not happened. b) It does not consider the long-term effects, such as brand affinity.

Even when using pragmatic approaches benchmarks should be used, targets should be set before the campaign takes place, and some effort should be made to consider the counter-factual.

Using market research to evaluate social

Social media tends to be widely dispersed and often very transient; as a result most conventional market research tracking studies do not have a big enough sample to effectively detect and measure it. However, traditional market research often has a valuable contribution to make – particularly in the context of long-term benefits, such as intention to buy, likelihood of recommending, and affinity.

Measuring the impact of social media via native metrics gives an indication of the strength and reach from the perspective of the campaign. Conversely, traditional research (e.g. tracking studies) tends to be based on a representative or relevant target group and can measure the phenomena from the orientation of the audience.

Key challenges for market research include:

- The sample size may not be large enough to identify impacts among the audience.
- Measurements often depend on recall, which can be unreliable.
- Challenges in terms of identifying exposed versus not-exposed.

Key advantages include:

- Focusing on the target group, as opposed to be reached by the campaign.
- The ability to measure multiple channels, especially when passive data collection is included in the mix.
- The ability to measure longer term issues, such as changes in intention, beliefs, and attitudes.
- Being able to measure people who have not interacted socially with the campaign, i.e. have not clicked, played, shared etc.

Over recent years there has been a growth in the ability to create research panels which combine the passive tracking of activities (including exposure to advertising and marketing) with surveys, and this is addressing some of the challenges faced by using market research to evaluate social.

In addition to using traditional market research to evaluate social campaigns and activities market research has developed a new/additional approach, social media research, which is covered in the next chapter.

Social Media Research

Social media research is a growing field and can be used to assess marketing campaigns and activities. . The use of social media research includes assessing campaigns in traditional media, in social media, and the evaluation of integrated campaigns and activities.

Measuring the effectiveness of communication campaigns through traditional media such as TV advertising has long been the remit of quantitative researchers across the globe. Representative sample surveys aimed at measuring the public's awareness of a campaign, recall of its messages and more importantly whether it has shifted the needle are the norm. The advent of social media, and the unprompted brand mentions it can yield, offers an additional opportunity to get a read on the effectiveness of campaigns and activities.

Social media research is a relatively new discipline within market research and one that offers a vast range of uses, including ideation, qualitative research, PR evaluation, and customer satisfaction. However, this Guide focuses solely on the use of social media research to measure and evaluate marketing campaigns and activities.

The sorts of questions that social media research is used for include:

- Has a TV campaign led to people talking about the brand in a different way?
- Are people posting messages about interactions with a new outdoor advertising campaign?
- Have people noticed a change in service?
- Are people engaged or cynical about a new social media campaign?

Definitions

Social media research tends to be described in two ways, using a narrow and a broad definition.

- **Narrow definition**. Quantitative research based on locating, collecting, and analysing naturally occurring discourses within social media. Examples include counting mentions of particular terms, measuring the sentiment of posts on a particular topic, and monitoring the volume of key phrases.

- **Broad definition**. Qualitative and quantitative research that either use naturally occurring social media discourses as their source material or socially created research situations. Examples of this broader category of social media research include: netnography, insight communities, and research about social media usage, as well everything covered by the narrow definition of social media research.

More information about social media research is available from The Handbook of Online and Social Media Research.

This Guide focus is on the narrow definition of social media research, i.e. the quantitative evaluation and measurement of campaigns and activities through the collection of naturally occurring discourses in social media.

Using social media to evaluate campaigns

Whilst the details of using social media research to evaluate campaigns can be complex the basic steps are fairly straightforward to describe:

1. Collect mentions from social media, for example from Twitter, blogs, Facebook (from public locations) etc. Creating a corpus, i.e. a database of mentions, tagged with date and other metadata. The data is often collected via third-party services, such as Brandwatch or Radian 6.
2. Clean the corpus of irrelevant mentions. For example, if evaluating the brand Apple, mentions of fruit are likely to be irrelevant. Remove or separate mentions from non-target

groups. Examples of non-target groups might be people from other countries or mentions from PR companies.

3. Evaluate the volume and content of the mentions. The volume tends to be simply the number of mentions, or the number of people commenting. The content of mentions can be measured in terms of sentiment (positive, negative, and neutral), or a wider/deeper level of analysis can be applied.

In social media research there is limited meaning to absolute values, therefore most researchers prefer to monitor several brands or campaigns, in order to benchmark the brand/campaign of interest.

The key areas of discussion in terms of measuring brands and campaigns via social media are:

a. Whether to use automated content/sentiment analysis, manual analysis, or some combination of the two.
b. The extent to which mentions in social media are representative of the wider population.

Strengths and weaknesses of social media research

One of the key strengths of social media is its immediacy, making it an excellent way to get an early read on what people think of campaigns, potentially within the first hours of its launch. The fact that posts are self-generated and can be mined retrospectively is also a key asset for social media as a data source.

Social media research means that researchers do not have to rely on respondents' recall, as with more traditional methods, and can potentially measure unprompted awareness from the level of mentions the campaign receives in social media. It also means that benchmarks of awareness and perceptions prior to the campaign can sometimes be derived after the campaign has ended, as there are fewer time constraints. This is a key advantage that traditional research does not have. Although, misuse of this phenomenon can lead to fishing rather than measuring.

Social media can reveal which aspects of a campaign are the most salient to the public in their own words and contexts. In traditional research, respondents are prompted by the research process, which may frame or change their responses. It has been argued that social media research provides a purer reflection of consumer perceptions and attitudes towards campaigns and activities and ultimately how they affect brand image/beliefs/attitudes than those derived through traditional research techniques.

Social media does not just enable measurement; it can also provide an in-depth understanding of initial reactions to a campaign, which could only be replicated through qualitative research techniques. This is because social can represent unprompted, unpremeditated, and unmediated responses from people going about their everyday lives.

Social media research is often perceived as a low cost option, which it can be sometimes. However, the use of the better software platforms and the amount of analyst time required can make social media research more expensive that it might initially seem.

One key concern is about social media's representativeness (or some would say lack thereof) of the public's opinion. Despite the fact that the reach of social media is expanding daily and that in the UK Facebook has a reported active user base of over 31 million and Twitter 15 million, the demographic representativeness of this audience is sometimes questioned. A more nuanced concern is that not everybody who uses social media uses it to post comments about brands and services – which could result in some voices being given too much importance.

Some researchers argue that as long as the lack of representativeness is recognised when contextualising and interpreting the content of conversations, it is a secondary issue. These concerns reinforce the need for social media not to be used in isolation from other data collection techniques

to provide context. The key question is to assess whether the attitudes and perceptions expressed in social media conversations reflect those of a wider audience.

The increasing use of hashtags by brands which serve as prompts to the campaign can to some extent remove the candid nature of social media conversations about these activities, generating something akin to prompted mentions. This should be considered when analysing results and analysed separately if appropriate.

Using social media research in conjunction with other approaches and metrics

As well as being used on its own, social media research can be used in conjunction with other approaches. For example, in the evaluation section of the appendix there is an example from Bottom-Line Analytics that shows how social media scores can be used in conjunction with metrics from digital, mass media, and retail to conduct media mix modelling.

Social media advice

As with other aspects of this Guide, the term 'Best Practice' has been avoided as there is not yet a settled and validated consensus on the best way to conduct this form of research. However, the advice presented here is gleaned from the leaders in the field.

- Run a benchmark analysis prior to the campaign. This step is key to measuring any shifts in levels of conversation about the brand, but also existing attitudes and perceptions. This will also be a useful exercise to determine which metrics the campaign should utilise. Using a three month time frame before the campaign is likely to smooth out most of the spikes driven by other events, activities, or campaigns.

- Build an intelligent search query. Using the campaign strapline or title will not be enough to gather relevant content. Use key words which relate to key elements of the campaign, for example central characters and premises and also key words associated with the themes or topic covered. This will ensure that the range of content gathered is in consumers' own words.

- Remember that volumes and share of voice can hide rich insights. While volumetrics are sometimes useful, they are not the be-all and end-all of social media analysis. The key to understanding the effectiveness of social is about measuring and not counting. This is why human analysis is often important in this context.

- Decide on the balance of human and automated analysis. There is a broad consensus that fully automated solutions are not yet capable of producing the required level of sensitivity. Analysis that is entirely based on human analysis requires sampling and can be expensive and lengthy. The most attractive options tend to combine human analysis with automation, often via some form of machine learning/training.

- Where appropriate, explore what research and data is available directly from the platforms.

Sentiment analysis

Sentiment analysis, in particular automated sentiment analysis, is a cornerstone of utilising social media to evaluate marketing campaigns and activities. However, it is a cornerstone that attracts a large amount of concern and indeed criticism. Sentiment analysis has attracted criticisms in terms of the accuracy of the sentiment analysis, the representativeness of online comments, and lack of established models that translate online sentiment to specific business outcomes.

What is sentiment analysis?

Sentiment analysis relates to a wide variety of approaches that seek to analyse text (i.e. qualitative information) and to ascribe quantitative values to it. For example, a sentiment analysis system might take a large number online comments about a brand and try to determine the proportion that were

negative, positive, and neutral (but it can also be used to ascribe more complex codes). In terms of campaigns and activities the aim of sentiment analysis is often to monitor performance over time.

The typical process for sentiment analysis is:

1. Collect a body of text, called a corpus.

2. Clean the corpus. Cleaning the data can include steps such as removing references to irrelevant brands (for Coke this would include removing references relating to the drug coke), possibly removing duplicates, and filtering to restrict the text to comments from an area or from a specific group (for example from consumers but not from marketing partners).

3. Apply analysis to code or score the text, for example into positive, negative, and neutral.

Whilst sentiment analysis can be conducted manually, either on a whole data set, or on a sample of a data set, most of the interest in sentiment analysis, in the area of social media, is about automated sentiment analysis.

Sentiment analysis – a two-dimensional approach

A campaign should be measured on two dimensions, the strength of the response to the campaign (typically the volume of mentions) and the sentiment of the response to the campaign. The strength of the response to the campaign can be further refined to look at: how many people made comments, how many comments each person made, and how influential are the people who made comments.

The queries about sentiment analysis tend to stem from the measurement of the sentiment, rather than the measurement of volume.

The Different Approaches to Sentiment Analysis

There are a variety of approaches to sentiment analysis. The key variations being:

- **Manual coding**: In manual coding a human coder, or a team of people, review the statements and assign codes to them. The code can be as simple as positive, negative, and neutral, but it can also be more complex. Unless the number of statements/comments is small, manual coding requires a sampling approach. When sampling statements/comments, care must be taken to sample them in a way that does not introduce bias.
- **Fully automated coding**: In fully automated coding all of the work is done by the software. The software can utilise a wide range of approaches. For example, some are based on natural language processing, some code on key words, some apply content analysis, whilst others are more dictionary/database driven.
- **Machine Learning**: Machine learning approaches are a hybrid of manual coding and automated coding. In a typical case the process starts with some of the comments being coded manually (for example perhaps one thousand comments). The next step is for the software to evaluate the manual coding, allowing it to create rules and then to code the rest of the corpus. Some systems have a process to identify text that is 'hard to code', allowing human coders to deal with these cases, which in turn improves the learning process.
- **Mechanical Turk**: MTurk (Mechanical Turk) is a crowdsourced service offered by Amazon. Using MTurk is a form of manual coding, but one where the costs tend to be lower, and the level of oversight is less. Using MTurk is often accompanied by screening for skills and by coding the same item more than once to increase the reliability.

Whilst no system is perfect, the general consensus is:

1. Manual coding is the most meaningful, but is less consistent than automated coding, and is usually slow and expensive.

2. Fully automated coding has a few fans and many critics.
3. Machine learning is an approach that tends to balance the criticisms and costs best, and is used by the case studies quoted in this Guide.

Increasing complexity of sentiment analysis systems

The earliest systems of automated sentiment analysis were based on identifying and scoring key words, for example scoring "love" +1, and "hate" -1. As systems became more complex the next step was to score combinations of words, for example "in the pink" as good, "in the red" as bad.

The next step was systems that decomposed phrases and attached them to specific objects. For example the phrase "I hate coffee but I love Coke" identified that the post had negative sentiment for coffee and positive for Coke.

The next level of complexity comprises a wide range of approaches, most of which are designed for purposes much more complex than simply scoring sentiment; their main purpose tends to be text analytics, automated services, and translation. The key ones are:

- NLP (natural language processing) seeks to interpret the meaning of the text, applying rules that have been generated from machine learning. Typical uses include translation and writing summaries.
- Statistical scoring of phrases, for example the way that Google Translate operates. Probability values are assigned to patterns that are observed.
- Content analysis seeks to decompose the text into codes, which are grouped into concepts and from there an overarching structure can be created.
- Discourse analysis seeks to understand what people do when they make meaning and employs approaches such as conversation analysis, psycholinguistics, and stance-shift analysis to work out what the people making the posts and comments were seeking to do.

The rate of development in this field is rapid and new options regularly appear on the market.

Issues surrounding sentiment analysis

Automated sentiment analysis of social media discourses arrived on the research and marketing scene about ten years ago with a rush of hyperbole, claiming that it was going to fundamentally change the way that consumers' voices and views would be assessed and measured. However, time and experience has identified a number of problems and limitations with social media listening in general and sentiment analysis in particular.

How representative are the comments?

Not everybody uses social media, and not everybody who uses social media posts comments. Consequently, people have queried how representative online comments are. This is an evolving discussion and will not be resolved soon. However, whilst concerns about representatively should be taken into account so should the ability of sentiment analysis and social media research to reach places that other research can't reach.

How good could sentiment analysis be?

The upper limit of how 'good' machine coding can be is determined by the agreement that can be achieved by human coders, which in turn determined by the complexity of the tasks. In a typical campaign evaluation, the inter-coder agreement is typically about 80%, so it is hard to envisage of machine coding ever being better than 80%, in these sorts of cases. If the machine was 100% in agreement with one of these two hypothetical coders, then it would only be 80% in agreement with the other.

However, machine coding is much faster and much cheaper than using human coders, and it can be much more consistent.

In general, the consensus is that if machine coding were to be 70% accurate then it would provide a really useful input into the evaluation of campaigns and brands.

Do we mean positive or 'on message'

Simple sentiment analysis tends to classify text as positive or negative. However, for some campaigns that is not an appropriate classification. Dr Stuart Shulman has been mining Twitter to assess an American anti-smoking campaign. The ads are shocking, showing people who have had surgery after cancer, so most of the people for whom the campaign would appear to be working are using words that would be associated with negative sentiment. Shulman's analysis of these campaigns tended to identify who noticed the campaigns, and whether the post exhibited fear arousal (for example "*Man that smoking commercial where the lady has a hole in her throat scares me every time*") or fear rejection (for example "*That smokefree commercial is bull shit. My grandma has been smoking cigarettes since she was 10 and she doesnt have a hole in her damn neck*").

It is likely that as sentiment analysis develops it will move beyond positive and negative, as a number of systems already do.

Other uses of sentiment analysis

Sentiment analysis can be used for much more than the analysis of marketing campaigns and activities, for example:

- **Customer satisfaction**. Social media monitoring and the analysis of sentiment can be used to assess customer satisfaction.

- **Service delivery**. Sentiment analysis can be used to assess the reactions to service users and facilitate real-time management of services, from traffic management, to retail, to users of social services. One of the key ways that modern customers request support or express anger is to shout it out in social media, which allows brands to listen and act.

- **Competitor analysis**. Social media discourses, and the sentiment they reveal, can be used to assess competitors as well as an organisation's own brands. This is useful in terms of competitive intelligence and essential in bench-marking the discourses relating to an organisation's brands.

- **Reaction marketing**. Analysing social media, and assessing the sentiment of those discourses, can allow an organisation to react to opportunities. The form of reaction might be to offer help, to promote a solution, or to create a campaign.

- **Commercial information**. Thompson Reuters has incorporated Twitter sentiment analysis into its Eikon trading platform - 'the first mainstream financial platform to provide Twitter sentiment'. This follows an SEC decision to allow companies to communicate news to investors on Twitter.

Sentiment analysis guidelines

Key advice includes:

- Sentiment analysis is more valuable in combination with other approaches.
- The results for the target brand, campaign, or activity should be benchmarked against other brands, campaigns, or activities.
- The corpus should be checked and cleaned.
- Determine whether the analysis is to be completed on the whole data set or on a sample – manual analysis usually requires sampling.
- Manually check some of the statements assigned as positive and some assigned as negative, to gauge the level of accuracy/confidence.

The merging of research, service provision, and marketing

Social media research is one of the research areas at the front of the move towards a more integrated picture of marketing and market research. A social media listening tool might be used simultaneously to:

1. Track sentiment.
2. Identify individuals with problems and offer solutions.
3. Identify sales interest at the individual level.

The ethics and regulatory framework for this research and marketing integration is not clear or complete and is likely to evolve over the next few years.

The Future

This final section the main body of the Guide is a collection of views from leading voices in the area. Each expressing hopes, fears, aspirations, and expectations.

Stan Sthanunathan

Senior Vice President, Consumer & Market Insights, Unilever

There is little doubt that Social is already having a big impact on our business. It is one of a handful of major, inter-connected forces that are fundamentally reshaping both marketing and marketing research practices. Social provides us with new, powerful ways to connect with the people who buy and love our brands and deepen their experience. But it also changes the competitive dynamic. So brands such as Chobani, in the USA,can quickly build sizable market presence cost effectively by understanding demand and creating word-of-mouth.

Yet, we believe that the impact of Social on research has only just begun. The potential is certainly significant. It is effectively a new form of ethnography that allows us to understand the people who most care about our brands and the issues that touch the categories they operate in. Done well it can be just-in-time, truly scalable research. However, currently few of the many tools available have the rigour to fulfil this potential.

In time, Social will not only shape 'what' we do as research professionals, but 'how' we do it. Market research and marketing will progressively converge as data collection becomes part of marketing and interpretation of numerous, often messy and ambiguous data streams, becomes more central to adaptive planning. This convergence across data streams, analytics and decision tools will require a profoundly new model of how we work with Agencies.

Above all, we must commit to shaping the future not just following trends. Understanding that Social drives business value will, of course, be critical. But we must be wary of hiding behind the demands for demonstrable ROI for innovative approaches if it forces us to be followers not shapers of the future. In this rapidly evolving field, the need is to have a perpetual beta mind set combined with a heavy dose of curiosity and inclination to take risks. Remember: If you do the same type of research and expect deeper insights, it is the definition of the height of optimism.

Rory Sutherland
Vice Chairman, Ogilvy Group

One of the great problems caused by the explosion of new forms of media is that people have sought to understand this highly complex new environment through the lens of communication rather than through the lens of human psychology.

The distinction is important. People do not absorb messages - they process them, evaluate them, interpret them and draw probabilistic inferences from them before deciding whether or not to act. There are very good evolutionary reasons for this: any naive individual who took all information at face value would have exited the gene-pool fairly rapidly, as a victim of deception by others.

Any evaluation of a medium simply on measures of how many people it reaches is woefully incomplete. You need to know how people process communication in that medium. And you need to understand that the inferences they draw may have almost nothing to do with the ostensible content

of the message: for instance a wedding invitation sent by email will contain exactly the same information as an engraved wedding invitation sent on a gilt-edged card, but your expectations of the wedding will be dramatically different either way. Our information processing routinely goes into 2nd and 3rd order considerations of intent.

One of my favourite examples of this inference comes from social media: in reading hotel recommendations on rating sites, people were heavily influenced not only by the content of the comments, but by the standard of grammar and spelling in which comments were written. They used the information not only to infer whether people liked the hotel, but also to make assumptions about the kind of people who stayed there.

The future of utilising social needs to take the lens of human psychology into account.

Edwina Dunn
Co-founder of dunnhumby and CEO of fan science company Starcount

For decades, data has been the key to unlocking secrets about people: their preferences, behaviour and habits.

What have changed are the parameters by which people can be understood through this data. When we started out the best data was from the Census, so it was a case of, 'you are where you live.' Then with shopping basket data it was 'you are what you eat'. Now the most important source is social media and the paradigm that interests us is 'you are who you follow'.

Social media has so much potential to deliver a new level of insight, because you are dealing with brands which people care passionately about. This isn't about what people buy but what they love and how they spend their spare time. Similarly, the reward for loyalty is not financial, but recognition: being retweeted by a global celebrity or star.

At Starcount, we estimate that approximately 160,000 stars account for 60% of total traffic. They are the fuel behind social networks: the brands and stars that people are interested in.

The future for social media is to find new ways to connect fans and followers to the people and things that they are interested in. People hate being bombarded with information and offers that they don't need or want. Relevance remains the most important asset in customer relations, and the most successful organisations will be those who understand this and tailor their services to each individual's unique needs and interests.

Jake Steadman
@jakesteadman, Head of Research EMEA, Twitter
Social media measurement is evolving in three key ways:

Integration
In the past, social media measurement often used different metrics to the broader business. CMOs would see reports that evaluated campaigns across traditional media using consistent KPIs combined with platform-specific figures for the social elements.

The tools available to measure social media campaigns have evolved and teams have started measuring success using more standardised

KPIs. Measuring Retweets or Likes is being replaced by assessing how effectively content has driven brand metrics like Awareness, Consideration, or Purchase Intent. This trend is accelerating, media teams are deploying more sophisticated (yet often more traditional) measurement tools with social channels. Social media measurement will increasingly be fully aligned with standardised brand and commercial KPIs, ending platform-specific reporting.

Redefining Value

Social media is much more than a marketing channel. Social can fundamentally re-shape a brand's relationship with customers, introducing more transparency and interaction. Social can help generate new customers and improve the service offered to existing customers. We are working with a growing number of brands to quantify the value that best-in-class customer service on Twitter provides for their business, in terms of cost efficiency and brand attraction.

There is a need now to measure everything that a social business represents. For example, the impact of great customer service, the value of additional trust rising from transparency, the impact of empowering staff to Tweet and act as brand advocate. These will be actively planned, delivered, and measured as social media becomes more pervasive across organisations.

Commercialising Research

In the past, research teams were often separate from the commercial operation of a business. They looked after traditional research, but responsibility for analysing social campaigns often sat with separate teams. As researchers have embraced social media as a source of insight and installed new platforms to analyse it, they've started to become the barometer of a brand's social media presence.

This is a huge opportunity for research teams to more closely align themselves further with the commercial side of their businesses and to advise on how best to adapt to a social-media world.

Research teams have the tools to measure the value of social media, the neutral position in the business to report back without bias, and the authority to present their findings as the best solution.

??? ????

LinkedIn

Beyond the popularity contest - A bigger view of social media measurement

Think of social media metrics and your mind instantly jumps to Shares, Likes and Retweets. Now it's true that these numbers can have plenty of value. They are a potential indicator of engagement levels for brand content, and an increasingly important distribution channel for that content. At LinkedIn we see the power of recommendation and social proof playing out all the time when it comes to amplifying the impact of the content that brands deliver. But we also know that this is only part of the picture.

Our experience of watching members consume and engage with content proves that there's a vast amount of untapped potential out there where social media measurement is concerned. That potential involves moving beyond a focus on what brands say, and starting to listen more meaningfully to the content created by audiences themselves. At the simplest level, this can include following the thread of discussions in relevant groups, but the potential for scaling up social media listening and integrating it with your campaign planning and measurement is huge. This was the reason why we launched LinkedIn Trending Content, which aggregates a view of the subjects that different LinkedIn audiences are engaging with, at the same time as Content Marketing Score, which quantifies how much engagement a brand's own content is delivering. The value of social media

measurement isn't just in showing that your content is popular – it's in ensuring that your content is actively addressing the issues that your audience cares about.

Ray Poynter

Author: **The Handbook of Online and Social Media Research & The Handbook of Mobile Market Research**

In producing this guide I have had the chance to see a wide range of alternative predictions for the future, a process which has shaped my views of the benefits and challenges. The key benefits that I expect to see are:

- *Ever larger amounts of data, in terms of coverage and granularity, combining media experience, social activity, and behaviour.*
- *The growth in new sources of data, such as location-based research and the use of wearables.*
- *Better text analytics, allowing open-ended comments to become even more central to interpreting what people say and think.*
- *The development of bots to interact with people in social, to evaluate what they are doing and to supply services and responses.*
- *A shift to measuring the real world, as it is lived, rather than via constructed instruments and tests.*

The two key challenges, for the evaluation of social, are going to be:

1. *The challenge of separating correlation from causality will become harder with more data, from more sources, collected from real life, rather than from carefully designed exercises.*
2. *The challenge of remembering that the biggest brand effects are long-term, in a world where the short-term can be measured so effectively and the results of the short-term used to manage campaigns and activities.*

The big change, as far as market researchers are concerned, will be the increasing merging of market research and marketing.

Coda

The five key takeaways from this project are:

1. Measurement should be designed in from the start, not added retrospectively. This means having a clear strategy, fitting the method to the budget, objectives, and context.
2. Measurement should take both short-term and long-term effects into account. Focusing only on short-term goals can make the long-term benefits harder to achieve.
3. Earned media tends to be the icing on the cake. Rewards are more often dependent on paid and owned media, working in integration with other media. When earned media does deliver significant results it is typically not inexpensive.
4. New models of influence, targeting, and homophily are challenging long-standing assumptions of how marketing works, generally reducing the importance of influence and targeting and increasing the importance of homophily.
5. Integrated campaigns make it very hard to be certain about attribution and causality, which means assumptions have to be made, stated, and factored into the evaluation.

Appendices

The appendices include:

- Case studies and examples
- Formal evaluation of campaigns
- The history and context of influence
- List of contributors
- Glossary
- References

Case Studies and Examples

The #IPASocialWorks project has collected a number of case studies which have been presented as examples of social being used for effective commercial return. The project has examples for all of the hypotheses listed above and many of them, of course, are demonstrations of more than one hypothesis.

Whilst some of the cases have stronger measurement methodologies behind them, all of them have tried to measure the impact of their activity and have some value.

Note, we are looking for more case studies, so please contact us if you have interesting examples.

The case studies included here are:

- ASB Bank
- BT Retail
- Department for Tourism, Philippines
- Doritos – Mariachi
- 'Dumb ways to die'
- Infiniti Europe
- Kern's Gnomes
- Kerry Foods / Mattessons – access via IPA website
- Miss Universe – NBC
- O2
- Starbucks Tweet-a-Coffee
- TfL – access via IPA website
- Visit Iceland – access via IPA website

ASB Bank

ASB are a New Zealand Bank whose share of the home loans market was under pressure from competitors driving down the rate in a very commoditised market. The bank wanted to challenge the market and chose social activity to do this. Their objective was to convert their social media activity from being a complementary service channel to a revenue generator, by building a database of warm home loan leads to convert. ASB created a product called the "Like Loan". This campaign won a Gold award from WARC in 2014.

The principle of the 2013 campaign is that as more visitors visited their product app on Facebook on a given day, the bank lowered the rate. One winner would then be offered the rate the number of likes 'created'. This was repeated over four weeks. The banks reports that the campaign created 17,778 leads, of which 11,555 were new customers. At the point where ASB submitted the campaign to WARC it had already converted 229 of the leads into loans.

The campaign cost was reported as NZD 378K including loss of 'interest' revenue, but the value from the conversions to date is NZD 4.5 million. The campaign was run again in 2014, adding TV to the mix.

http://awards.digitalstaging.co.nz/likeloan

BT

Since 2009 BT have been developing a multi-faceted social approach, including:

- A social monitoring tool (Debatescape) to listen to what customers are saying about their service experience and the brand.
- A customer service team to proactively and reactively offer help to customers complaining in the social web.
- Embedding social media activity within call centres using trained existing technical help advisors.
- Setting up social media presences in the places where conversations were taking place, and using Twitter to enhance communications, and creating opportunities for self-hep and community support.

BT divide their metrics into Soft (e.g.: Facebook likes, followers on Twitter, positive mentions, Views on Youtube etc.) and Hard (e.g.: Customer Satisfaction, Customer Effort, Likelihood to churn, and Call Deflection).

The calls deflected, for example, show impressive returns. Over the time period measured (12 months) Twitter deflected 38,023 calls – saving £98K ; YouTube deflected 55,000 calls – saving £72K and the Forums deflected 504,000 – saving £942K. The ROI model has been built on volume of contacts, unique customers, effectiveness (full resolution of query) and cost of contact. In 2013 BT stated that the lower effort, lower cost, and positive brand impact have resulted in a £2 million saving.

In terms of satisfaction scores, BT have revealed detail of the initial impact after the inclusion of using social media to offer customer service. A comparison of 3 Month Rolling Average Satisfaction scores revealed that in October 2010 the score was 63% and by Oct 2011 this had improved to 87% (a 37% increase).

Importantly, they also found that Customer Effort was the main driver of the BT Net Promoter Score. They asked the question "How easy was it to get the help you wanted from BT today?"

See the full case study on the IPA website.

Department for Tourism, Philippines

The Philippines had an image problem that prevented potential tourists from appreciating the positive appeal of visiting the country. Annual tourist arrivals had flat lined at below four million, one of the worst performances in the region. In 2012, a campaign was developed to drive tourism numbers and to make the Philippines, and the campaign, the most talked about, shared and 'searched for' in Southeast Asia.

To make up for the lack of marketing funding, the campaign relied on crowdsourcing the creative work taken from an initial idea: "It's more fun in the Philippines". This captured the idea that the Filipino people enhanced the experience of any tourist. And, through harnessing the power of social media, 95 million Filipinos were transformed into a massive and creative tourism salesforce.

The campaign became the No.1 global trend on Twitter just two days after its launch, with the slogan mentioned online once every minute. The Philippines achieved a 231% increase in Google hits for the country, while tourism arrivals hit an all-time high of 4.3 million an increase of 9% exceeding the target of 7% (source: Philippine DoT). The campaign brought in more additional visitors than tourism powerhouse Malaysia, despite having a far smaller marketing budget. It is asserted that the campaign was amplified by social media and over 60,000 suggestions for "It's more fun in the Philippines" were created across the world. By making it easy for Filipinos to grasp the campaign and take part in it, these thousands of memes were inspired by user-created 'ad-maker' sites, and iPhone, Android and Windows apps. Individual ideas covered the entire spectrum of subject areas in

the tourism promotion master plan, later forming the basis of promotional posters, print ads and even a global TV commercial – entirely from crowd-sourced ideas.

Interestingly, the tourism levels did not slump after a major typhoon hit the islands in November, despite widespread global media coverage, but continued to rise and is targeted to reach another an all-time high in 2014. The video case from the Spikes Asia award entry : http://www.spikes.asia/winners/2013/effectiveness/entry.cfm?entryid=2006

Doritos – Mariachi

In the last 5 years, Doritos had targeted a youthful audience of 18-24 'socially active snack consumers' and the brand has grown its value to become the 3rd biggest in the UK crisp and snacks category (Nielsen). The brand decided it was time to focus on consumers outside of its heartland for the first time – on the 25-44 year olds, and created a business objective of increasing value sales by more than 6%.

Doritos had been positioned as the perfect snack for social 'get-together's with friends. However, the new target was to create a role for Doritos in any social occasion, with any consumer audience. Given Doritos is a tactile, informal snack they felt they could position the brand as a way of breaking down formality and awkwardness. They expressed this as the brand proposition: *Doritos breaks down formality and gets you in the mood to party*. Their solution came in cheesy pop cover songs. Research revealed these songs had universal appeal. The idea was to create a cheesy cover's band - 'Mariachi Doritos' – that captured the spirit and fun of the Doritos brand, and acted as a metaphor for the role Doritos could play at any party to break down formality and help people let their hair down.

The band appeared at social occasions up and down the country in real people's living rooms, real people's dinner parties, birthday parties, and BBQ's. The aim was that customers experience the brand as part of an "*all year round*" campaign.

High-performing' content was pushed harder through all social media channels. The criteria was that if content performed well within 3 hours (past a 2% virality benchmark) then it would be amplified to wider audience. Driven to the campaign-hub from other social channels (including Twitter and YouTube) as well as from TV, fans and non-fans of Doritos could go to the Facebook page and sign up to 'join the tour'. While TV played a role, its role was to drive awareness and interest in the social media tour, and the belief was that this approach would be much more cost-effective than a traditional TV-only approach. This is an interesting reverse of the usual format.

As well as engaging consumers, they also created Mariachi Doritos content tailored to Doritos retail partners as an entertaining way of thanking them for their support and encouraging them to give Doritos future preferential display in store.

http://www.youtube.com/watch?v=mx46IwhxxXg

OCRs (Nielsen Online Campaign Ratings) and Nielsen Brand Effect were used to measure the performance of the social media content and for the first time to help PepsiCo understand the impact of social media on purchase intent, brand love etc. 'Soft' brand measures moved positively, including purchase intent and recommendation intent, and sales for the brand increased by 11%. At the time of writing about this case, no modelling or other data has been supplied to suggest any attribution by channel for the sales increase.

'Dumb ways to die'

The 'Dumb ways to die' campaign was created in 2012 for the transport authority in Melbourne, Australia to raise awareness of preventable deaths, and has won multiple awards. The campaign utilised paid media (newspapers, outdoor advertising, and local radio), owned media (website and mobile app), and earned media (Tumblr and YouTube). The YouTube video has been viewed over 72 million times (http://youtu.be/IJNR2EpS0jw).

The campaign elements can be assessed separately, for example how many people have viewed and shared the elements. But the key issue that the campaign sought to address was safety, and there the picture is more complex, at outlined in The Age, 15 September, 2013 (http://www.theage.com.au/victoria/safety-message-may-be-falling-off-20130914-2trrr.html). The key points were:

- There had been a rise in suicides and near misses before the campaign, which was why the campaign was commissioned.

- Following the launch of the campaign (November-January 2012-13) there was a 30% drop in near misses.

- But, the first six months of 2013 saw the highest number of near misses ever - and 9% more than the same period in 2012.

The figures for near misses have led to discussions about a) campaign wear out and b) what would have happened if the campaign had not been run – i.e. the issue of the counterfactual.

Infiniti Europe

Study published on the IAB UK website http://www.iabuk.net/sites/default/files/research-docs/IAB%20measurement%20framework%20for%20social%20media_FINAL4_0.pdf .This case is being actively considered by the IPA Peer Review Group.

- **Intent**: "Reach out and empower key opinion formers to talk passionately about Infiniti coming to Europe so that our prospects get to hear about us"

- **Activity**: 4 CGI 40 second films shared with key influential bloggers via SMR (Social Media Reputation) and uploaded to multiple video sharing sites in Europe.

- **The plan**: In terms of the IAB's four As, the campaign targets were 50% Appreciation, 20% Awareness, 20% Action, and 10% Advocacy.

The chosen KPIs are selected in the table below:

Awareness	Appreciation	Action	Advocacy
No of blog posts	No of plays	Click through rates (CTR)	No of emails
Search rankings	No of comments	Website referrals	No of embeds
	Cost per engagement (CPE)	Cost per lead (CPL)	

The results of the campaign included:

- Assets ranking in top 3 for vehicle searches in video sharing sites such as YouTube and Metcalf.

- 893K plays of the videos.

- Every €100 invested there were 848 plays

- €0.11 CPE (cost per engagement)

- Every €100 invested there were 24 web referrals

- €4 CPL (cost per lead)
- Click through rate nearly twice industry average
- Lower than average advocacy metrics

Kern's Gnomes

Kern makes precision scales for the science and education sectors and they were able to utilise a social media campaign to grow awareness, SEO ranking, sales, and business leads. The campaign by Ogilvy was based on creating a crowdsourced science experiment, where people with precision scales in different parts of the world measured the weight of a gnome (from Kern) to help create a global map of the impact of the shape and the mass of the earth on the apparent weight of the gnome.

The campaign, in 2012, utilised a blog, sets of scales and gnomes in a special postable package (allowing the experiment to be forwarded), and social media promotion. People were encouraged to request a kit to conduct the experiment, and to upload images and information. Within two weeks it has been claimed that the campaign reached 355 million people, across 152 countries, with thousands of websites linking to the GnomeExperiment.com website. The experiment has been featured on BBC and Discovery Channel and has been the subject of a TED talk.

The claimed return on the campaign includes, within two weeks of the campaign:

- 22% increase in sales to the target group.
- 2,200 former customers getting back in touch.
- 1,445 new leads.
- Number 1 research on Google searches for precision scales.

This IPA Review Group is currently looking for additional information about this case.

More information about this example is available at
http://www.cipr.co.uk/content/events-awards/excellence-awards/results/12-global-public-relations-campaign/case-study

Miss Universe – NBC

An interesting example in the world of broadcasting occurred in the last broadcast of 'Miss Universe'. 160 countries voted live for their favourite through NBC, which meant 23 million votes in 9 minutes. Facebook and Twitter buttons were at the bottom of a second screen application on phones, tablets or desktops. There was a pre-loaded message on the bottom of the graphics which enabled viewers to send a voting message to their followers, for example, 'I have just voted for X'. At the end of the show analytics showed that 12% of the players on the second screen who were playing the game and watching the show, had come from the social share button.

Indeed the area of amplification is increasingly important as brands look not only for social to perform specific functions by themselves, but also to magnify other activity e.g. on TV.

O2

O2 is a leading digital communications company in the UK, with over 23 million customers. The brand realised that social media data was a valuable source of consumer insight, and that social channels would be key to cost effective customer service.

O2 partnered with social intelligence agency Face to develop their own platform which they called RTO2 (Real-Time O2). RTO2 was launched in 2010 across the O2 UK business with over 400 users.

RTO2 enables O2 to monitor and instantly react to messages directed at the brand on the social media channels where it has a presence (Twitter, Facebook, and YouTube). The platform collects the messages and augments them with 'metadata' to add contextual information which is extremely

useful in customer service decision making. For example, the system collects: sentiment, customer influence level, customer location, customer history, service area, and type of issue.

Using social data in this way enabled the newly formed Social Media Response team to prioritise messages by topic, negativity or influence level and ensure their reply takes into account all previous social interactions with that customer. RTO2 provides a real time reporting dashboard, with the following metrics: Response time, % Issues resolved, Team member activity, Incoming and outgoing messages, Messages by channel, and Messages by service area.

In July 2012 O2 suffered its biggest marketing and communications crisis ever – its network went down for over 10 million customers for a two day period. Naturally, customers took to social media to share their concerns, frustrations, and anger. Over 30 team members were assigned to respond to the surge in mentions. The ability to react to every mention as it appeared online significantly impacted sentiment surrounding the issue and overall brand perception. Because O2 could manage much of the servicing of customers during this difficult period in-house, using their social media systems rather than outsourcing the problem, their analytics system meant they could determine the amount of money saved per customer.

O2 use their tool to analyse the reach of conversations arising from Marketing and PR campaigns, their competitive position in the social media landscape, and understanding the role of influencers, particularly when targeting specific content areas. Influence is measured in three main ways :

• Who is talking about O2 the most?(Volume);
• Whose mentions reach the biggest audience? (Visibility);
• Whose mentions drive the biggest reactions from other people? (Engagement/Influence).

Influencers are also used for promotional offers, for example a retweet competition for new handsets.

O2 has calculated how much each point of customer satisfaction is worth in net profit per customer. Using this tool, they have started to establish how much social media interaction is worth. Certainly, social customer service is now an important part of O2 brand identity. But they have also provided the following figures to help to demonstrate its value: Customers who have interacted with the organisation via social media deliver a customer satisfaction score of 73 whereas those who do not, deliver 69 – i.e. social media interaction is associated with customer satisfaction scores that are 6 percentage points higher.

Starbucks Tweet-a-Coffee

This campaign, launched in October 2013 illustrates how close the link between marketing and selling can be in social media. The Tweet-a-Coffee campaign allowed people to buy their friends a Starbucks coffee by tweeting the friends Twitter handle and "@tweetacoffee". The process results in a $5 gift card for the recipient. Research firm keyhole tracked the use of @tweetacoffee and estimated that the campaign generated $180,000 in sales with more than 27,000 people taking part. In addition to awareness Keyhole estimate that Starbuck's have linked Twitter IDs, mobile phone numbers, and customer IDs for about 54,000 people.

This IPA Review Group is currently looking for additional information about this case.

More information available from Mashable.

The Formal Evaluation of Campaigns

This section looks at the three main, formal methods of evaluation campaigns:

- Market mix modelling
- Experiments and A/B testing
- Event or individual/household level modelling

Market Mix Modelling

Market mix modelling (MMM) has been the conventional 'gold standard' approach to measuring communication effectiveness for some years. Typically, it seeks to isolate the impact of advertising and other marketing impacts through a statistical analysis of aggregate weekly sales and marketing data. Its key advantage is that it is a flexible technique that can quantify quite complex relationships in the context of the overall marketing mix.

As with any technique, MMM has limitations when measuring conventional marketing, but there are additional challenges with social:

- Social media (and digital media more generally) is hard to measure consistently and so capture in a model.
- The impact can be relatively small (compared to price and promotions), so is lost amongst the background noise.
- Social activity is often to reinforce the perceptions of brand 'loyalists', so the impact is longer-term.
- It works in combination with other media along the consumer journey, so needs to be measured 'synergistically' with other media.
- It is not always clear what is causing what. Great marketing might lead to online conversation and buzz, which is indicative of sales but not predictive. This can, in particular, be a problem with weekly or monthly aggregated sales, where there is insufficient time granularity to separate which effect comes first.

Model design is very important when using MMM to measure social. For example, the Nielsen's Digital Media Consortium has suggested that poorly designed marketing mix models can understate the ROI from Facebook by as much as 48% and from Google search by up to 39% (Neff, 2014). However, the same study also found that 25% of the value ascribed to paid search should actually be shared with TV, print, or other digital media. One reasons for these sorts of effects is that conventional MMM typically measures short to medium-term response in sales, but advertising often works over the longer term by reinforces existing habits etc. If long-term effects are being sought then the MMM needs to be appropriately configured (Cain, 2008).

Data Inputs

The following data input issues need to be addressed.

- Issues such as cross-media comparability, synergy, and granularity should be addressed.
- Interactions between social and other media need to be modelled.
- An attempt needs to be made to assess causality, for example via Granger Causality. Nielsen's analysis of how Twitter leads to larger programme audiences is a good example of using Granger causality (Nielsen, 2013). The study found that Twitter *caused* significant changes in live viewing for 29% of episodes tested in the USA.
- Consider whether to use VAR, (Vector AutoRegression).
- Consumer journey modelling builds a model of the consumer journey as a series of links, with functions linking each step, for example paid search might be a function of TV and earned media, and sales a function of paid search.

MMM Data Inputs - Best Practices

When organising the data for a MMM project there are a number of key considerations and broadly accepted best practices including the four points below.

Cross-media comparability. Metrics should be comparable to other channels and provide a measure of exposure. These should be based on both reach and impressions to guard against 'excessive' frequency. Engagement metrics (clicks, likes, shares, comments, number of followers/fans, etc.) tend to be a poor measure of brand activity, are not closely linked to sales (e.g. Nielsen, ComScore, DataLogix research), and potentially suffer from bias in interpretation because the causality can be in the 'reverse direction'.

Granularity. Models are best built at the most disaggregated geographic level possible (i.e. weekly or finer), to maximise the chance of measuring small differences e.g. TV region, postal sector etc.

Synergy. Model connections and feedback inherent in and between social and other digital media, do not just model direct impacts (e.g. TV + search + social).

Media channel granularity. Model *Paid*, *Organic,* and *Viral* streams separately where possible. Similarly, model PC social separately from mobile.

Modelling interactions

A more fundamental issue with traditional models is that there is a complex pattern of interaction between the media and sales variables. For example, TV helps drive search, which in turn drives social exposure. Similarly, social helps build the brand, but, also, stronger brands will have more people who are likely to engage with the brands social activity. This means there is a feedback loop from social to sales and from sales to social. In the parlance of econometrics, social exposure and engagement isn't exogenous (determined outside of the system) but largely endogenous (determined by other marketing variables).

There are several different techniques to improve conventional MMM and to overcome the issue of 'endogeneity'. It is important to remember that there is no single 'best' approach. As with all models, the choices and decisions largely depend on the most important features of the model and the data.

Granger Causality

At the heart of the issue of measuring media linkage and synergy is how to measure 'causality'. The simplest situation is when we want to know whether a social metric can be a causal indicator for sales or another behavioural KPI. This might be part of the model checking process or an issue in its own right. For example, Twitter is often used as a lead indicator of TV ratings, but does Twitter 'cause' higher viewing or is it 'caused' by programme popularity.

The standard technique for establishing causality is based on a test originated by Clive Granger, and known as 'Granger Causality'. Granger causality is based on 'predictive causality' i.e. if A causes B, then knowing A will help make a better prediction of B. More formally, a variable X is said to Granger-cause Y if knowing about the history of X and Y allows a better prediction of Y, compared with only knowing the history of Y.

The concept seems straightforward, but is typically only part of an analysis. Granger causality can help identify whether e.g. Twitter leads to higher TV ratings, but it doesn't quantify how big the impact is. It is directional, so it is usually used as a step before a more formal model is built. One key point is that Granger causality is based on past values - it does not help determine causality between two events that happen at the same time.

One problem to look out for is the *post hoc ergo propter hoc* fallacy (after this, so because of this) – which describes the problem of thinking that if A keeps follow B it is caused by B. Two variables can look like they are inter-related but are actually determined by a third. This can be a particular issue

in social and viral marketing because of homophily – i.e. that similar people group together. For example, a product launch can look like it is driven by viral or word of mouth marketing because sales adoption is clustered. However, the 'cause' can be broadcast media. This can be a particular problem if there is time aggregation bias e.g. analysis is done at aggregated weekly level, but the 'time-ordering' happens faster. For example, social buzz responds to a promotion and sales respond to promotion.

Vector AutoRegression (VAR)

A more comprehensive method of accounting for interactions between media variables and feedback from sales is based on vector autoregression (VAR). VAR is an econometric methodology which seeks to measure interdependence between time series variables. Its key strength, according to its supporters, is that it accounts for complex interactions between variables in a relatively 'theory free' way, i.e. the analyst is not imposing a structure on the data, or deeming some variables to be exogenous. It has been popular in macroeconomics for some time and is now being applied in marketing.

Besides being able to capture interactions between marketing variables, VARs can to some extent capture long-term dynamics. Given that the economics of advertising frequently rests on its long-term impact, being able to trace the impact on base sales is important. Conventional MMM will typically use adstocks as a measure of the long-term impact on base sales, where VARs are unconstrained and can detect the impact of any marketing driver that influences habit.

While VARs have many benefits in principle, they face a number of practical criticisms:

1. Prioritising data mining over 'theory' puts a great emphasis on data quality and data collection. The impact of measurement error or missing variables can be amplified through the models. And, if too many variables and time slices are added then the model will be 'overfitted', producing a great description of past data, but at the cost of less accurate about the future.

2. Measuring causality in social is very difficult. VARs often use weekly data - and this can be insufficient to tease out whether A is causing B, or B is causing A, for example, whether earned media generated sales interest or sales stimulated online conversation among the brand's user base.

3. VARs use deep correlations in the data to measure long-term marketing impacts. This can prove problematic for measuring social when there is a long-term trend to using social media (driven by technology) and a long-term trend in sales (for example, driven by marketing or the economy).

An example of VAR in use is a study reported by Stacey et al (2014) which study found that 62% of store traffic for an electronic retailer was driven by paid marketing, and the predominant channel was achieved through generating consumer conversations about the products. Within this mix, Twitter was effective at helping to stimulate conversations. Another good example of using VAR to measure interactions is a study looking at Twitter & TV (Rubart, 2013). A Twitter-sponsored study for UK mobile operators found that Twitter improved the efficiency of TV advertising by 35%.

Consumer Journey Modelling

Granger causality and vector autoregressions have deep traditions in economics and time series econometrics, but are often not chosen by practitioners because the approaches can magnify data errors through amplified feedback loops (so leading to overestimated impacts of media ROI). The simplest alternative approach is to model elements of the consumer journey directly in a linked and chained series of models. For example, modelling paid search as a function of TV and earned media and then including the search terms in a sales model.

Modelling systems of equations has a long tradition in econometrics. An important requirement with this approach is to make sure that variables that are determined simultaneously are treated as endogenous. For example, if search leads to more social exposure, but social leads to more search, then the models should be treated as simultaneous. The extent of simultaneity often depends on time aggregation: there is more unobserved feedback in weekly than in daily data.

There are a number of techniques for measuring systems of equations that are used by practitioners besides conventional simultaneous modelling (e.g. Two Stage least squares). For example, Structural Equation Modelling is useful when there is a latent (a hidden or unobserved) variable that has an impact on sales. For example, 'word of mouth' is not observable but is partly driven by earned digital media. Another option is to use Bayesian Networks to model interrelationships. One of the advantages of Bayesian approaches is that they can provide more control of interrelationships, so they can be more stable when there is noisy and incomplete data.

For more information on consumer journey modelling see Peter Cain (2014a & b).

Experiments and A/B testing

Marketing mix modelling (MMM) is the conventional method of measuring advertising effectiveness for above the line media. However, one of the benefits of most digital media is that it is possible to construct experiments, and online brand tracking studies have made use of this approach for some time.

An increasingly common way to measure social media ROI is to fuse or link social data to household panel or purchase data. For example, both Facebook and Twitter have worked with DataLogix to trace online exposure to offline sales. Data is typically either linked through cookies, email, and/or other identity measures, and then analysed anonymously.

There are two elements to experimental studies:

1) Controlling the exposure.
2) Measuring the effect (see Kohavi, 2008)

In terms of the exposure, the aim is to control it as the individual level, options that allow people to have multiple accounts, such as cookie tracking, can be flawed as they can mean some double-counting.

Controlling exposure

The gold standard approach to testing/experiments is randomised controlled trials, where people are randomly exposed to media and the only difference between the groups is their exposure. For example, serving the test ad or survey in random rotation.

While some academic studies may achieve this high bar, true random experiments are expensive and are not routine. In particular, many 'social' campaigns are not random by design. So, while it is possible to create randomised exposure to paid social advertising (within a specific channel), exposure to the social/earned element is not random (i.e. the interactions between consumers). Importantly, the consumer-to-consumer interactions are a key element of social effectiveness.

When it is not possible or economic to create control exposure groups, the exposed/non-exposed groups are created after the activity has run. This technique is called matching. Matching is widely used in some social sciences (e.g. politics) and is also known as propensity matching. The idea is to create a quasi-matched sample based on observable characteristics of the sample. For example, after the activity has run, the data provider might create two samples from the data set, each with identical demographics, social media usage, etc. The only difference should be that one group was exposed to the advertising, and one wasn't.

Whilst it appears easy to simply compare site A vs site B, or treatment A vs treatment B, the research should try to control for the counterfactual – i.e. what would have happened anyway, for example what would have happened if neither A or B had occurred.

Measuring experimental effects

There are lots of potentially confounding factors that can frustrate the accuracy of the analysis (time of day, site, offline media). Without careful control, each element can significantly bias the results. Social commerce, for example, might provide an easier way for brand loyal consumers to purchase the brand - but many would have done so anyway, so the incremental sales are not as high as they might appear.

With experimental approaches it is still important to measure uplift, not just increased propensity, i.e. the difference in likelihood of responding given being in the treatment group versus being in the control group (mathematically, P(O = 1 | x; Treatment) - P(O = 1 | x; Control)). Quite often, approaches just measure propensity through a binary response model (e.g. logit, probit). [1]

Event or individual/household level modelling

Inevitably, there are a variety of techniques that mix elements of the modelling and experimental approaches. Typically, they use consumer or respondent level data. In the context of social, there is a strong advantage to some of these approaches because they can capture or model the non-random way that social messages diffuse through the population.

Three key approaches, consumer mix modelling, attribution models, and a combination of consumer mix and experiments should be considered.

Attribution models

'Attribution modelling' is a term that is applied to many different techniques but is typically used to refer to how digital media work together across the digital consumer journey.[2] Most approaches are based on modelling converter versus non-converter paths using sophisticated analytical techniques (e.g. machine learning, Shapley value etc.)

A good example of how attribution modelling can be used to measure social media is provided by attribution modelling specialist Visual IQ (2012). The study finds that Facebook is undervalued by traditional methods (typically, by around a third) because it helps drive demand at the 'top of the funnel'. This finding is reasonably consistent from other attribution models.

---

[1] Bias can be further mitigated using difference in difference estimators, i.e. looking how changes between the treatment and the control group change over time.

[2] In digital analytics, attribution modelling has a wide range of meaning. For example,
- Fractional attribution, the credit for an action (e.g. download or sale) is divided across multiple factors, according to a fixed ratio. - Multichannel attribution, these models typically apply to multiple digital channels, and require that each channel employ tagging/tracking - Algorithmic attribution, weights are given to different factors in a dynamic process, based on an algorithm, usually proprietary. - Last click, a single source approach, based on the last click before the action - Last non-direct action, this the basis of the standard Google Analytics conversion - First interaction, the first thing that happened gets all the credit - Linear attribution, all the steps in the path get equal weight, by definition this can't be as good as the right model, but will be better than most bad models - Time decay model, the first step gets a low value, the second interaction more, through to most for the last step.

Consumer Mix Modelling

Consumer mix modelling is econometric modelling at the consumer level, typically using sales panel data. The advantage of modelling at the individual respondent level is the ability to exploit variation in exposure across the data but controlling for exposure variables (similar to matching), which means that social can be measured in the context of price, promotions and other marketing instruments. This also means that communication can be measured against consumer behavioural segments (brand loyals, trialists, etc) and not only capture ROI, but also how the advertising is working.

Consumer mix modelling relies on good single source data or data links, for example, the GfK Media Efficiency Panel (see Cadbury/GfK 2010). Its weakness is usually that the media exposure data, particularly for above-the-line media, is not single source and is estimated/fused.

Combination of consumer mix and experiments

The most sophisticated approaches combine experiments and very advanced analytical techniques. This allows the study to control and understand the interplay between viral transmission (i.e. non-random exposure), homophily (i.e. non-causal response), and message. Currently, these approaches are used in academia but are not yet available at scale in the commercial sector.

A good example of these sophisticated analyses and the level of granularity of the findings is Sinan Aral's paper on the role of influentials and susceptible individuals (Aral, 2011). The paper uses an experiment and a model to unpick how a social content (a movie rating app) spreads in a social network. Of interest is the finding that typical estimates would over-estimate contagion by 300-700% because they don't account for similar people responding in the same way to external factors. In the study over 50% of the effect was homophily.

Example of MMM incorporating social media research

Bottom-Line Analytics supplied the project team with an example looking at a specific category of personal care. The inputs to the media mix model included: digital (including display, paid search, mobile, and social media ads), traditional media (including outdoor, radio, TV, print, and cinema), retail (including distribution and price), a factor to account for periodic seasonality, and two metrics based on measuring open ended comments in social media.

The social media measures and applied a proprietary approach utilising stance-shift analysis to generate proprietary scores for the target brand, topics and competitors. The study looked at weekly UK data from January 2012 to April 2014.

The scored social media engagement metric showed a strong association with retail sales volume, i.e. when consumer discourse around the brand was engaged and positive it tended to be followed by an increase in sales and when commentary was less positive, lower sales followed shortly afterwards. The study suggested that social commentary acted as a leading indicator.

Finally, supporting analysis using Structural Equation Modelling showed how a language based consumer engagement metric operated as a lower funnel metric. By contrast, digital display media residing on social media platforms were found to be an upper-funnel metric, with only an indirect impact via total brand awareness.

Further reading and resources

The documents listed below provide additional information and link to the points above about evaluation methods.

- The power of like 2, a report produced by Facebook and comScore looking at the impact of social media marketing, with a focus on North America. The_Power_of_Like_2-1.pdf The Power of Like Europe, a report by comScore and Facebook that develops the Power of Like 1 and 2 studies and focuses on Europe. comScore Power of Like Europe - REPORT.pdf

- Understanding Paid and Earned Reach on Facebook
- Generating business results on Facebook

Advice when conducting formal evaluation of campaigns

The following four points should serve when undertaking the formal evaluation of campaigns, ofr example when using MMM or A/B testing.

1. Use more than one measurement technique, seeking to measure the entire consumer journey. No single approach is perfect. Combinations of modelling and experimental methods work well, particularly as part of a test and learn process.

2. Biases in aggregate modelling or A/B testing approaches can be severe, resulting in spurious findings. Always be sceptical about very high ROIs unless there is a clear consumer-driven reason. Check the scale of the impact makes sense with supporting bottom-up calculations.

3. With all modelling approaches, use granular data to avoid social signals being drowned out by other elements of the marketing mix.

4. If using aggregate sales data, model social as part of a 'system'. Think first of the consumer journey, then introduce more sophisticated statistical approaches where these are warranted.

The History and Context for Influence

This section provides a briefing on influence, looking at the history of influence, some cases studies, the key players, current debates, and further references.

Key References and Timeline

This section shows how today's thinking on the subject has come about, drawing attention to key thinkers and publications.

Katz, E & Lazarsfield, P, 1955, Personal Influence: The part played by people in the flow of mass communications, The Free Press.

This book grew out of Lazarsfield's study of the 1940 US Presidential election (The People's Choice) and produced the Multistep Flow Model. Lazarsfield and the Multistep Flow Model suggest that influence flows from the media to opinion leaders through to the majority, i.e. the majority are not being influenced directly by the mass media.

Milgram, S, 1967, The Small World Project, Psychology Today

Stanley Milgram was responsible for the famous six degrees of separation experiment, termed the small world project. Packets were distributed to randomly selected people in Omaha and Wichita, USA, who were requested to send them to people they did not know in Boston, Massachusetts. Since people did not know the recipients they forwarded the packages to people who might have a better chance of getting it to the right person. Most of the packages were never delivered, but amongst those that were delivered, the average number of steps was close to six – generating the popular idea that on average, everybody in the US is connected by six degrees of separation.

In 2000, Malcom Gladwell drew on this study to suggest that hubs were important to the flow of influence. Gladwell highlighted that Milgram found that about half of the successful packets went through the hands of just three people.

Cialdini, R, 1984, Influence: The Psychology of Persuasion: Issues for social marketing and social evaluation, William Morrow.

Robert Cialdini looked at the psychological mechanics of how things flow from one person to another and produced his six principles of influence:

1. Reciprocity: if people receive something that is a 'favour' they may feel obligated to return the favour.

2. Commitment: once people choose something they are inclined to stick with it.

3. Social proof: sometimes called safety in numbers, doing what others are doing.

4. Liking: people are more likely to be influenced by people they like.

5. Authority: people tend to follow ('obey') people they see as being in authority, e.g. uniforms, titles, experts etc.

6. Scarcity: people are more inclined to want things that they think have become scarce, particularly if they have become scarce because they are popular.

Several of these principles reappear in later writings by others, for example Social Proof is close to the topic covered in Herd by Mark Earls, and Authority underpins the mavens that appear in Feick & Price, Gladwell, and Berry & Keller.

Feick, L & Price, L, 1987, The Market Maven: A diffuser of marketplace information, Journal of Marketing

Feick and Price introduced the term Maven, or market maven, to the marketing literature, as somebody who seeks knowledge about products and services and who is seen as knowledgeable and

is therefore influential. This study underpins many of the books and papers advocating influence, and is very widely cited.

Gladwell, M, 2000, *The Tipping Point: How little things can make a big difference, Little Brown*
Gladwell brought the concept of influence and social copying to the general public. Gladwell, proposed that diffusion is based on three types of influencers: Connectors, Mavens, and Salesmen. Gladwell populates his book with a large number of examples of how ideas, fads, trends, and messages have travelled through communities, starting with the humble Hush Puppy, through to Paul Revere's ride through the night in April 1775 summoning support to fight off the British.

Gladwell highlights the small world experiment of Stanley Milgram and notes that about half the successful packages travelled via just three friends of the target person. Gladwell declared that these three were the social hubs, and that these sorts of hubs are essential to the success of ideas being diffused through networks. This proposition underpins much of the thinking about influentials, and is robustly attacked by writers such as Duncan Watts.

Keller, E & Berry, J, 2003, *The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy, The Free Press*
Ed Keller & Jon Berry's subtitle sets out their position clearly, a small group influence behaviour and the rest of us follow. Keller and Berry's definition of influentials is typical of the market place and is grounded in common sense. "*the Influentials are active in their communities. They are highly engaged in the workplace and in their personal lives as well. They are interested in many subjects and are connected to many groups. They know how to express themselves and do so.*". This definition overlaps with Gladwell's three hub types, the Salesmen, Connectors, and Mavens.

The Keller and Berry concept of influence has the advantage that a population can be screened to find the influentials, for research or marketing.

Reichheld, F, 2003, *One Number You Need to Grow*, Harvard Business Review

Fred Reichheld's Net Promoter Score is based on measuring people's propensity to say they would recommend a product or service. This measure taps into interest about word of mouth, advocacy, and influence. The NPS route does not assume that some people are more influential than others, it aggregates the NPS score across the relevant user/customer base, in this way it is closer the thinking of Duncan Watts, which downplays the idea that special people are actually special.

Watts, D, 2004, *Six Degrees: The Science of a Connected Age, WW Norton*
Duncan Watts brought Stanley Milgram's six degrees of separation experiment up to date, collecting over 50,000 email chains from over 160 countries, producing a more nuanced picture than Milgram's experiment. Watts' experiment suggested that six links between strangers was about right, which in turn might suggest it is a feature of networks themselves. However, Watts' findings led him to refute Gladwell's focus on influencers/hubs. Firstly, Watts' experiment (which was much larger than Milgram's experiment) did not produce hubs of any great size (5% being about the top, as opposed to about 50% in Milgram's experiment). Secondly, Watts showed that the results of this sort of experiment can be ascribed to network theory as opposed to individuals with influence.

Watts has shown that a simulation of 10,000 artificial consumers, with a few simple rules and a few influencers, can replicate the behaviour of influencer models – with fewer assumptions.

Watts says the critical thing is whether people are ready to be persuaded, rather than whether somebody is persuasive. In Watts own words "*If society is ready to embrace a trend, almost anyone can start one--and if it isn't, then almost no one can,*" (Thompson, 2008). This may well link to another explanation of shared behaviour, homophily, which suggests that people in networks connect with similar people, and their similarity is evidenced in the way they do similar things.

Earls, M, 2007, HERD: How to change mass behaviour by harnessing our true nature, Wiley
Mark Earls pioneered, in the marketing world, the concept that the most important thing about people's behaviour (including what they buy) is that "*Our species is first and foremost a social one.*" This view of the social leads Earls to conclude that "*Word of mouth is the most powerful sales tool.*"

On the topic of influencers, HERD is more equivocal, reporting people like Keller and Gladwell, but doubting that influence is primarily about 'special' people. However, it concludes "*If you have to prioritize some customers over other, seek out those who have the most influence over their peers (either in terms of the number of people they influence or the degree or manner in which they influence others).*"

Bentley A and Earls M 2008: Forget Influentials: copying is how brands spread, Admap October
This paper argues that we have misunderstood influence as a push-force (something that one individual does to another) rather than a pull-force (based on copying). The paper argues that influence is not what is said or heard but more importantly what is observed. The paper asserts the empirical evidence does not support the assumption that all (or even most) human social networks are hub-and-spoke shaped (as per the influentials hypothesis), nor are they fixed (as the network analogy tends to suggest).

Vogelaar, R, 2011, The SuperPromoter: The power of enthusiasm, Palgrave
Rijn Vogelaar has promoted the concept of identifying, working with, and amplifying people he terms superpromoters, people who love your brand and who tend to 'good mouth' it. Vogelaar asserts that focusing on detractors can be very wasteful, since it rarely turns them into promoters and takes attention away from people who like or even love the brand.

Bentley, A, et al, 2011, I'll Have What She's Having: Mapping Social Behavior (Simplicity: Design, Technology, Business, Life), MIT Press
Alex Bentley, Mark Earls, Michael O'Brien, and John Maeda explore social behaviour and the degree to which relatively simple models explain social and market phenomena. Drawing on the work of Duncan Watts and others the authors show that complex rules and guiding hands are not required to explain a wide range of market behaviour. Their conclusion is that people are massively influenced by what others do, but that the role of influencers is much less. They comment that celebrities have some impact "*but their influence on each of us is much less than we imagine.*"

The two key points that the authors make, in terms of influencers, is that we should not think of social networks as being connected by wires, it is much more fluid and less solid than that, and that the process is much less predictable than many seem to assume, it only looks predictable with hindsight.

Solis, B, 2012, The Rise of Digital Influence: A 'how-to' guide for businesses to spark desirable effects and outcomes through social media influence, Altimeter Group
Brian Solis asserts that social influence metrics link to social capital, which he terms the opportunity to be influential. Solis identifies three pillars which help explain social capital:

1. Reach, broadly how many people are connected with somebody.
2. Relevance, the fit between the topic and the person.
3. Resonance, a combination of quantity and strength of the person's output.

The report then presents a method of leveraging influencers and shows case study examples, from Microsoft, Virgin America, and Starbucks.

Keller, E & Fay, B, 2012, The Face-to-Face Book: Why real relationships rule in a digital marketplace, InkWell Publishing
Keller & Fay seek to draw a major distinction between digital social and social, pointing that that most conversations about brands take place offline, with face-to-face being the most frequently

used medium. The book uses as an example the big bet Pepsi took in 2010, when it moved most of its media spend to social, which was followed by a loss of market share, a loss of share of throat, and by 2011 a resumption of traditional media as its lead media. For example, it returned to advertising during the Superbowl after just one year's absence.

The book highlights a study by MarketShare in 2011 which looked at a variety of clients and which concluded "*Social voice represents a critical pathway through which more than half the impact of paid advertising and media passes in generating consumer purchases.*"

Whilst advocating the power and utility of WOM and influence Keller and Fay are clearly cautious about some areas, not only in highlighting the role of face-to-face, but also in terms of things like influence scoring systems such as Klout. They comment "*the jury is still out on whether services like these actually provide a true measure of influence.*" and "*We expect that people ultimately will conclude online influence is one piece of the puzzle, but wholly insufficient on its own.*"

Aran, S, 2012, Identifying Influential and Susceptible Members of Social Networks, Nature.

Sinan Aran is perhaps the most widely respected academic researching and publishing in the area of influence, homophily, and networks. Aral's work with both observation and controlled experiments has produced results that appear to show that influence is less common in social networks and less powerful than has been commonly supposed. For example, one study showed that not allowing for homophily led to a 700% overestimation of influence.

Aran highlights that one of the key problems with many earlier studies, and with common sense, is the lack of an examination of the counter-factual. i.e. what would have happened anyway.

Flores, L, 2013, How to Measure Digital Marketing: Metrics for Assessing Impact and Designing Success, Palgrave Macmillan

Laurent Flores' recent book focuses on the issue of metrics for digital marketing and includes a substantial section on social media related issues.

On the topic of influence Flores is rather dismissive of outdated PR models (e.g. number of likes, tweets, automated sentiment) and systems that assign a single value for somebody's influence (i.e. personal measures such as Klout and Kred).

Flores focuses on evaluating channels that are essentially mini-media channels, for example bloggers with a substantial following. Flores advocates a relatively manual process of evaluating options. He looks at the context of the brand/service/topic and the relevance and resonance of the potential channel (a mix of subjective assessment of the quality of the material produced by the channel and the impact the channel appears to have in terms of readers/links/ranking/citations/shares etc). Against these two, Flores assesses the risk that the channel conveys. The system aims to recommend a channel that is useful, rather than simply saying it is influential. The system identifies that information from well-connected people spreads better, but this seems to be a network or homophily phenomenon, rather than an influence phenomenon.

Lutz Finger and Soumitra Dutta, 2014, Ask, Measure, Learn: Using Social Media Analytics to Understand and Influence Customer Behavior, O'Reilly Media

Finger and Dutta pull together many of the strands of the previous texts quoted here and cite a number of large scale studies. In general, they report some evidence of influence, typically from outside a network, but very limited evidence, and very little influence within networks.

Players in Influencer Marketing

Two interesting elements in the influencer marketing picture are companies providing access to influencer groups and companies providing influencer metrics.

Influencer Groups

Influencer groups are people who are supposed to have influence and who have signed up to some sort of programme to help promote products and services, typically in return for rewards. Examples include: Klout and its Perks programme, TapInfluence, and P&G's Tremor

There are many forms of these communities. If these communities work, in the sense of producing a bigger net impact (after allowing for any extra costs) on target metrics (e.g. product trial or sales), it is likely to be a combination of two effects a) these people being influencers, b) these people being motivated (often extrinsically) to spread the message.

One of the key warnings from Sinan Aral's work is that influence marketing needs to consider the counterfactual, in particular how many of the recruited influencers would have bought the product anyway.

Providers of influence Metrics

Over the last few years, several measures of social influence have been developed and promoted, such as:

- Klout

- Kred

- mBlast

- Peer Index

- TweetLevel

- Appinions

Some people divide influence metrics into two groups, personal and contextual. Personal measures assign a score to somebody and that score is used for all contexts, for example Kred. Contextual measures, such as Appinions, seek to measure influence within a topic, which means one individual can have different scores for different topics.

Key Debates

This section addresses the key issues and debates.

What is an influencer?

There are a wide range of definitions, but three common strands are:

1. Users/buyers of a product or service who tend to influence others, either by making recommendations or being seen to use/buy things.

2. Third parties, such as blogger, journalists, experts, analysts.

3. Members of influencer communities, people who have been signed-up by an organisation to act as the agents of influence marketing, for example people in Klout's PERKS programme.

There are of course other definitions; the WOMMA Influencer Handbook describes five categories: advocates, ambassadors, citizens, professional/occupational, and celebrity.

Is the notion of influence bogus?

As the review of the influence literature highlighted, there is a debate about the extent to which influence exists. Duncan Watts has challenged the very notion of influencers, asserting "A rare bunch of cool people just don't have that power. And when you test the way marketers say the world works, it falls apart. There's no there there." (Thompson, 2008)

Others believe in influence, for example PR firms such as Burson-Marsteller claim E-Fluentials can make or break a brand. According to MarketingVOX, an online marketing news journal, more than $1

billion is spent a year on word-of-mouth campaigns targeting Influentials, an amount growing at 36% a year, faster than any other part of marketing and advertising."

Sinan Aral' work suggests that influence exists, can be measured, but is much smaller than is commonly assumed, and often does not reflect a cost-efficient way of promoting a product or service.

Influence, influencers and influenced?

Is influence a push or a pull-phenomenon? Common sense tends to suggest that influence is about push, i.e.an influencer impacts other people, for example by persuading them. However, much of the data reported in this section suggest it may be largely pull. Pull works by people, when susceptible, copying behaviour from people around them. For more on this point see Bentley Earls O'Brien (2011) and Earls M (2012).

What is the value of social influence metrics?

Brian Solis of Altimeter makes the point that companies who use social influence metrics, such as Klout, are looking at influence backwards. They are looking at the scores, rather than what makes the scores (2012).

Also, danah boyd [3]highlights the observer effect in metrics like Klout. People who buy into the system seek to improve their score, making the scoring system less valid, such that many of the high scoring people are players in the game, not necessarily the most influential. Whilst this concern is valid when looking at these metrics in comparing all customers, these metrics are probably very relevant when looking at influencer groups, where playing the game, and playing it well, are core parts of the process.

Is influence largely opportunity to see?

If somebody is rated as influential because they have 100,000 people a month visiting their blog, or because 10,000 people typically see what they post on the Facebook page or profile, the key to their influence could be just the number of eyeballs they are able to deliver. This model would be very similar to traditional measures for conventional media.

---

[3] danah boyd has [expressed a preference](#) for her name to be lower case

List of Contributors

Project Team

The core project team was:

- Christopher Wellbelove, BT
- Christian Walsh, MRS
- Fran Cassidy, Owner
- James Devon
- Jed Hallam
- Mark Earls, author
- Ray Poynter (lead author), The Future Place
- Simeon Duckworth, Mindshare

The project would like to following people for their contributions to the project, in terms of copy, comments, and/or reviewing material.

- Alex Bentley, University of Bristol
- Alison Henderson, TfL
- Ben Tuttle, Insight Catalyst
- Celina Burnett, Ohal
- Ed Keller, Keller Fay Group
- Edwina Dunn, Starcount
- Francesco D'Orazio, FACE
- Gaelle Bertrand, Precise
- Helen Bryant, RSPB
- Jake Steadman, Twitter
- Jennifer Thompson, The Walt Disney Company
- Laurent Flores, INSEEC Business School
- Malcolm Bell, Zaggora.com
- Masood Akhtar, Bottom-Line Analytics
- Matthew Stockbridge, Mondelez
- Patrick (Paddy) Barwise, London Business School
- Patrick Bray, stream:20
- Paul Edwards
- Peter Cain, Marketscience Consulting
- Rory Sutherland, Ogilvy Group
- Stan Sthanunathan, Unilever
- Sue York, NewMR
- Tom Kerr, Tesco Bank

Note, the fact that people have made contributions to the project does not mean they necessarily endorse or agree with the entire report.

Glossary

Item	Description
A/B Testing	A/B Testing is a method of identifying which elements perform best. In a typical case, two randomly assigned cells are shown executions that differ and the results compared. A/B testing can be used to measure the effect of different treatments, or it can be used to select a best performing execution.
AVE	Advertising Value Equivalent, the cost of buying the space taken by a piece of 'earned' content, if it had been purchased. AVE has been used by the PR industry to assign a value to earned media, but it is generally seen to be an overestimate.
Attribution Models	Attribution models seek to determine which factors contributed to an outcome and to assign values to the weight for each element. Key terms and models include: **Single source attribution**, all the credit is given to a single factor, such as last click or first click **Fractional attribution**, the credit for an action (e.g. a download or sale) is divided across multiple factors, according to a fixed ratio. **Multichannel attribution** models typically apply to multiple digital channels, and require that each channel employ tagging/tracking **Algorithmic attribution**, weights are given to different factors in a dynamic process, based on an algorithm, usually proprietary. **Last click**: a single source approach, based on the last click before the action **Last non-direct action**, this the basis of the standard Google Analytics conversion **First interaction**, the first thing that happened gets all the credit **Linear attribution**, all the steps in the path get equal weight, by definition this can't be as good as the right model, but will be better than most bad models **Time decay model**, the first step gets a low value, the second interaction more, through to most for the last step **Customised allocations**, different parts of the pathway to the final clicks are given different weights depending on some exogenous factor (such as belief).
CPI	Cost per impression, to evaluate how widely something might have been seen.
CPE	Cost per engagement, where engagement might have been liked, commented on, shared etc.
CPL	Cost per lead, where sales leads are tracked back to campaign elements.
CPO	Cost Per Order, the cost of the marketing/activity divided by the number of orders.
CPR	Cost per referral, for example where shares are tracked, or the use of discount codes etc.
CTR	Click-Through Rate, the percentage of people who see something who click on it.
Granularity	Granularity refers to how fine the detail is. One key issue with MMM is the time granularity. If data is supplied as monthly aggregates then it will often be impossible to determine what caused what, because during the month there will have been multiple changes and multiple outcomes. Weekly data is better than monthly, but may not always be granular enough to identify different actions and consequences.
MMM	MMM, Market mix modelling, is an advanced statistical technique that takes multiple

	measures of marketing activity, such as advertising, and seeks to estimate the contribution of each element to the outcome (e.g. to sales). MMM is considered the 'gold standard' in terms of assessing the impact of marketing activities, such as advertising.
ROI	Return on Investment. At one level the term ROI relates to a generalised concept, expressing what a project delivers. More formally it is the value of the return divided by the cost of the investment – for example ROI% = (Net Profit/Investment)*100.
ROMI	Return on Marketing Investment, is a subset of ROI, linking the marketing investment and the measured return.
Share of Voice	In the ad business this tends to mean impressions for a brand divided by all impressions. In social media research it tends to mean mentions of a brand divided by all mentions of other relevant brands.

References

- Anderson, C, 'The End of Theory: The Data Deluge Makes the Scientific Method Obsolete', Wired, 2008, http://www.wired.com/science/discoveries/magazine/16-07/pb_theory
- Aral, S, Content and Causality in Influence Networks, 2011, http://videolectures.net/icwsm2011_aral_influence/
- Aral, S, Muchnika, L, & Sundararajana, A, 'Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks', 2009, http://www.pnas.org/content/106/51/21544.full.pdf
- Barwise, P & Farley, J, 'The State of Interactive Marketing in Seven Countries: Interactive Marketing Comes of Age', Journal of Interactive Marketing 19, 1 (Summer 2005), 67-80.
- Barwise, P & Meehan, S, 'The One Thing You Must Get Right When Building a Brand', Harvard Business Review, 2010, http://hbr.org/2010/12/the-one-thing-you-must-get-right-when-building-a-brand/ar/1
- Bughin, J, Doogan, J, & Vetvik, O, '', McKinsey & Company, 2010, http://www.mckinsey.com/insights/marketing_sales/a_new_way_to_measure_word-of-mouth_marketing
- Cain, P, 'Limitations of conventional market mix modelling', Admap, April pp 48-51, 2008
- Cain, P, 'Advanced Methods in Marketing Econometrics', http://www.itqanalytics.com/community/blog/advanced-methods-in-marketing-econometrics#.U-H6iPldXe6, 2014a
- Cain, P, 'Marketing-mix Modelling Big Data', http://www.itqanalytics.com/community/blog/marketing-mix-modelling-big-data#.U-H6svldXe4, 2014b
- Carl, W, 'What's all the buzz about? Everyday communication and the relational basis of word-of-mouth and buzz marketing practices', Management Communication Quarterly, 2006
- Charman, S, 'Fake Reviews: Amazon's Rotten Core', Forbes, 2012, http://www.forbes.com/sites/suwcharmananderson/2012/08/28/fake-reviews-amazons-rotten-core/
- Corcoran, S, 'Defining Earned, Owned And Paid Media', Forrester, 2009, http://blogs.forrester.com/interactive_marketing/2009/12/defining-earned-owned-and-paid-media.html

- Court, D, Elzinga, D, Muder, S, & Vetvik, O, 'The Consumer Decision Journey', McKinsey & Company, 2009, http://www.mckinsey.com/insights/marketing_sales/the_consumer_decision_journey

- DeMers, J, 'How To Execute A Converged Media Content Strategy (And Why You Should)', Forbes, 2013, http://www.forbes.com/sites/jaysondemers/2013/09/11/how-to-design-and-execute-a-converged-media-content-strategy-and-why-you-should/

- Divol, R, 'Demystifying social media', McKinsey & Company, 2012, http://www.mckinsey.com/insights/marketing_sales/demystifying_social_media

- D'Onfro, J, 'Twitter admits 5% of its users are fake', Business Insider, 2013, http://www.businessinsider.com/5-of-twitter-monthly-active-users-are-fake-2013-10

- Farahat, A, & Bailey, M, 'How effective is targeted advertising', Yahoo, 2011, http://research.microsoft.com/en-us/um/beijing/events/ia2011/ayman%20sigir%20_mike_v2.pdf
- Finger, L & Dutta, S, 'Ask, Measure, Learn', O'Reilly Media, 2014
- Flores, L, 'How to measure digital marketing', Palgrave, 2013.
- Gelman, A, 'Causality and Statistical Learning', Columbia University, 2010 http://www.stat.columbia.edu/~gelman/research/published/causalreview4.pdf
- Harford, T, 'Big data: are we making a big mistake?', The Financial Times, 2014, http://www.ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz2yHpWXRpX
- Hodson, H, 'Google Flu Trends gets it wrong three years running', New Scientist, 2014, http://www.newscientist.com/article/dn25217-google-flu-trends-gets-it-wrong-three-years-running.html
- Keller, E, Fay, B, & Berry, J, 'Leading the conversation: Influencers' impact on word of mouth and the brand conversation', WOMMA, 2007
- Kohavi, R, Longbotham, R, Sommerfield, D, & Henne, R, 'Controlled experiments on the web:
- survey and practical guide', Springer Science+Business Media, 2008.
- Kornowski, L, 'Celebrity Sponsored Tweets: What The Stars Get Paid For Advertising In 140 Characters', Huffington Post, 2013, http://www.huffingtonpost.com/2013/05/30/celebrity-sponsored-tweets_n_3360562.html
- Lewis, R & Rao, J, 'On the Near Impossibility of Measuring the Returns to Advertising', 2013, Google and Microsoft, http://www.aeaweb.org/aea/2014conference/program/retrieve.php?pdfid=985
- Mutinga, D, Moorman, M, & Smit, E, 'Introducing COBRAs - Exploring motivations for brand-related social media use', International Journal of Advertising, 2011, http://brandprogramming.nl/wp-content/uploads/2011/04/Muntinga-Moorman-Smit-2011.pdf
- Neff, J, 'Digital Is Selling More Soap Than It Gets Credit For: Nielsen Study', Advertising Age, 2014, http://adage.com/article/digital/study-digital-return-investment-undervalued/291772/
- Ries, E, 'Why vanity metrics are dangerous', 2009, http://www.startuplessonslearned.com/2009/12/why-vanity-metrics-are-dangerous.html
- Rubart, C, 'New MarketShare Analysis Reveals Twitter Impact on TV Ads', MarketShare, 2013
- Silver, N, 'The Signal and the Noise - Why So Many Predictions Fail-but Some Don't', Penguin, 2012
- Solis, B, 'The Rise of Digital Influence: A "how-to" guide for businesses to spark desirable effects and outcomes through social media influence', Altimeter Group, 2012
- Stacey, E, Pauwels, H, & Lackman, A, 'Beyond Likes and Tweets: Marketing, Social Media Content, and Store Performance', The Center for Measurable Markting, 2013

- Thompson, C, 'Is the tipping point toast?', Fast Company, 2008
- Watts, D, 'Viral Marketing for the Real World', 2007, http://hbr.org/2007/05/viral-marketing-for-the-real-world/ar/1
- Cadbury/GfK, 'Marketing Mix Evaluator: Cadbury's Chocolate Charmer Campaign', 2010, http://www.scribd.com/doc/45468533/Cadbury-Campaign-Results-Dec-2010
- Nielsen, 'The Follow-Back: Understanding The Two-Way Causal Influence Between Twitter Activity And Tv Viewership', Nielsen, 2013, http://www.nielsen.com/us/en/insights/news/2013/the-follow-back--understanding-the-two-way-causal-influence-betw.html
- Visual IQ, 'A Visual IQ Perspective - Four Crucial Takeaways on Attribution & Facebook ROI', 2012, http://www.visualiq.com/sites/default/files/four-crucial-takeaways-on-attribution-and-facebook-roi.pdf