Chapter 15

Evaluating marketing communications

As part of the marketing communication process it is necessary to evaluate the overall impact and effect that a campaign has on a target audience. It needs to be reviewed in order that management can learn and better understand the impact of its communications and its audiences.

Aims and learning objectives

The aims of this chapter are to review the ways in which marketing communications activities can be evaluated.

The learning objectives of this chapter are to:

1. discuss the role of evaluation as part of marketing communications;
2. explore the value and methods of pre-testing and post-testing advertisements;
3. explain the main ideas behind different physiological measures of evaluation;
4. provide an insight into the way in which each of the tools of the communication mix can be evaluated;
5. consider other ways in which the effectiveness of marketing communications can be evaluated;
6. measure the fulfilment of brand promises;
7. consider some of the issues associated with evaluating the effectiveness of digital and online communications.

For an applied interpretation see Jill Brown's MiniCase entitled *Measuring communication effectiveness at the Salvation Army* at the end of this chapter.

Introduction

All organisations review and evaluate the performance of their various activities. Many undertake formal mechanisms, while others review in an informal, ad hoc manner, but the process of evaluation or reflection is a well-established management process. The objective is to monitor the often diverse activities of the organisation so that management can exercise control. It is through the process of review and evaluation that an organisation has the opportunity to learn and develop. In turn, this enables management to refine its competitive position and to provide for higher levels of customer satisfaction.

The use of marketing communications is a management activity, one that requires the use of rigorous research and testing procedures in addition to continual evaluation. This is necessary because planned communications involve a wide variety of stakeholders and have the potential to consume a vast amount of resources.

The evaluation of planned marketing communications consists of two distinct elements. The first element is concerned with the development and testing of individual messages. For example, a particular sales promotion (such as a sample pack) has individual characteristics that may or may not meet the objectives of a sales promotion event.

An advertising message has to achieve, among other things, a balance of emotion and information in order that the communication objectives and message strategy be achieved. To accomplish this, testing is required to ensure that the intended messages are encoded correctly and are capable of being decoded accurately by the target audience and the intended meaning is ascribed to the message. The second element concerns the overall impact and effect that a campaign has on a target audience once a communications plan has been released. This post-test factor is critical, as it will either confirm or reject management's judgement about the viability of its communications strategy. The way in which the individual components of the communications mix work together needs to be understood so that strengths can be capitalised on and developed and weaknesses negated.

> Testing is required to ensure that the intended messages are encoded correctly and are capable of being decoded accurately by the target audience.

Prediction and evaluation require information about options and alternatives. For example, did sales presentation approach A prove to be more effective than B and, if so, what would happen if A was used nationally? Predictably, the use of quantitative techniques is more prevalent with this set of reasons. This concluding chapter of Part 3 examines the testing and evaluation methods that are appropriate to all the tools of the communications mix and introduces ideas relevant to the measurement of online communications.

The role of evaluation in planned communications

The evaluation process is a key part of marketing communications. The findings and results of the evaluative process feed back into the next campaign and provide indicators and benchmarks for further management decisions. The primary role of evaluating the performance of a communications strategy is to ensure that the communications objectives have been met and that the strategy has been effective. The secondary role is to ensure that the strategy has been executed efficiently, that the full potential of the individual promotional tools has been extracted and that resources have been used economically.

Research activity is undertaken for two main reasons. The first is guidance and development and the second is prediction and evaluation (Staverley, 1993). Guidance takes the form of

Table 15.1   Four dimensions of IMC

Dimension of IMC	Explanation
Unified communications for consistent messages and images	Activities designed to create a clear, single position, in the target market, delivering a consistent message through multiple channels.
Differentiated communications to multiple customer groups	The need to create different marketing communications campaigns (and positions) targeted at different groups (in the target market) who are at different stages of the buying process. Sequential communication models based on the hierarchy of effects or attitude construct apply.
Database-centred communications	This dimension emphasises the need to generate behavioural responses through direct marketing activities created through information collected and stored in databases.
Relationship fostering communications for existing customers	The importance of retaining customers and developing long-term relationships is a critical element of marketing communications.

Source: Lee and Park (2007). Used with permission from WARC.

shaping future strategies as a result of past experiences. Development is important in the context of determining whether the communications worked as they were intended to.

The prevalence and acceptance of the integrated marketing communications concept (Chapter 9) suggests that its measurement should be a central aspect when evaluating marketing communications activities. One of the predominant issues surrounding the development of IMC is the difficulty and lack of empirical evidence concerning the measurement of this concept. In an attempt to resolve this Lee and Park (2007) provide one of the first multidimensional-scaled measures of IMC. Their model is based on four key dimensions drawn from the literature. These are set out in Table 15.1.

> One of the predominant issues surrounding the development of IMC is the difficulty and lack of empirical evidence concerning the measurement of this concept.

Each of these dimensions is regarded as separate yet integral elements of IMC. Lee and Park developed an 18-item scale, derived from the literature, to measure these dimensions. The use of this approach may advance our understanding of IMC and provide a substantial basis on which IMC activities can be measured. It is interesting to note that Lee and Park see IMC as a customer-only communication activity and choose to exclude other critical stakeholders from their measurement model.

Advertising

An IPA report in 1998 stated that 23 per cent of finance directors said that if business costs were under pressure they would cut marketing and advertising before anything else (Farrow, 1999). Among the reasons offered for this view was the feeling that advertising was extremely difficult to measure and thus problematic in terms of its overall contribution to the organisation. If in doubt, cut it. On a more optimistic note Fendwick (1996) suggests that there have been four main stages to the measurement of advertising effectiveness.

> There have been four main stages to the measurement of advertising effectiveness.

1. Direct response – coupon response.
2. Executions – measurement of consumer psychological responses to the way individual ads are executed. The recognition and recall techniques were developed and refined to reflect this approach.

3. Campaign evaluation (current age) – the evaluation of campaigns working over a period of time. The use of econometrics and modelling techniques to examine the influence of key variables typifies this approach.
4. Research nirvana (future age) – through the use of computers and vast data sets it will become possible to evaluate specific individual and panel data regarding various emotional and rational impacts of a variety of marketing communication messages. The effect will be to enable managers to adjust their communications messages and media quickly, efficiently and much more effectively.

The techniques used to evaluate advertising are by far the most documented and, in view of the relative sizes of the communication tools, it is not surprising that slightly more time is devoted to this tool. This is not to disregard or disrespect the contribution each of the communication tools can make to an integrated campaign. Indeed, it is the collective measure of success against the goals set at the outset that is the overriding imperative for measurement, as will be seen later.

Pre-testing

Advertisements can be researched prior to their release (pre-test) or after they have been released (post-test). Pre-tests, sometimes referred to as copy tests, have traditionally attracted more attention, stimulated a greater variety of methods and generated much controversy, in comparison with post-tests.

The effectiveness of *pre-testing*, that is the practice of showing unfinished commercials to selected groups of the target audience with a view to refining the commercial to improve effectiveness, is still subject to debate. Reid (2000) argues that pre-testing can be used positively to support campaign development, predictively to gauge likely audience response and generally to improve advertising performance.

The methods used to pre-test advertisements are based upon either qualitative or quantitative criteria. The most common methods used to pre-test advertisements are concept testing, focus groups, consumer juries, dummy vehicles, readability, theatre and physiological tests. Focus groups are the main qualitative method used and theatre or hall tests the main quantitative test. Each of these methods will be discussed later.

The primary purpose of testing advertisements during the developmental process is to ensure that the final creative work will meet the advertising objectives. It is better to help shape the way an advertising message is formed, rather like potters continuously review their progress as they craft their vases, than to make a pot and then decide that it is not big enough or that the handle is the wrong shape. The practical objectives of pre-testing creative work are to provide opportunities to optimise ads before publication and second, and if necessary, to terminate an advertisement before costs become so large and commitment too final. Changes to an advertisement that are made too late may be resisted partly because of the sunk costs and partly because of the political consequences that 'pulling' an advertisement might have. The Newspaper Marketing Agency (NMA) in conjunction with Millward Brown has pre-tested a large number of newspaper ads and has found that pre-testing newspaper ads can more than double ad recognition levels. See www.nmauk.co.uk.

> The primary purpose of testing advertisements during the developmental process is to ensure that the final creative work will meet the advertising objectives.

Once a series of advertisements has been roughed or developed so that its messages can be clearly understood, advertisers seek reassurance and guidance regarding which of the alternatives should be developed further. Concept tests, in-depth interviews, focus groups and consumer juries can be used to determine which of the proposed advertisements are the better ones by using ranking and prioritisation procedures. Of those selected, further testing can be used to reveal the extent to which the intended message is accurately decoded. These comprehension and reaction tests are designed to prevent inappropriate advertisements reaching the finished stage.

"To enjoy greater success with dogs, run ads in newspapers."

RICHARD WATSON, COMMERCIAL DIRECTOR.
NESTLÉ PURINA PETCARE

We needed to seriously grow our brand in a hugely competitive market and we'd not advertised in newspapers before. But after some valuable pre-testing research we produced a campaign that helped us to be top dog in the sector. To see for yourself why newspapers are dubbed 'The Attention Medium' just fetch the case study from the NMA website. It could be a marketing person's best friend.

NMA NEWSPAPER MARKETING AGENCY
www.nmauk.co.uk/advertisingeffectiveness

Exhibit 15.1 **Newspaper Marketing Agency**
A newspaper ad, placed in The Independent, promoting Nestlé Purina Petcare, the virtues of pre-testing and the NMA.
Created for The Newspaper Marketing Agency by VCCP Blue; copyright VCCP Ltd.

Pre-testing unfinished advertisements

Concept testing

The concept test is an integral part of the developmental stage of advertising strategy. The purpose is to reduce the number of alternative advertising ideas, to identify and build upon the good ideas and to reject those that the target audience feel are not suitable.

> The concept test is an integral part of the developmental stage of advertising strategy.

Concept testing can occur very early on in the development process, but is usually undertaken when the target audience can be presented with a rough outline or storyboard that represents the intended artwork and the messages to be used. There are varying degrees of sophistication associated with concept testing, from the use of simple cards with no illustrations to photomatics, which are films of individual photographs shot in sequence, and livematics, which are films very close to the intended finished message. Their use will reflect the size of the advertiser's budget, the completion date of the campaign and the needs of the creative team.

Concept testing, by definition, has to be undertaken in artificial surroundings, but the main way of eliciting the target's views is essentially qualitatively oriented, based on group discussion. This group discussion is referred to as a focus group and is a technique used by most agencies.

Once a client has approved the agency's plans, ad production can begin. Very often the creative team or an independent artist will produce roughs or drawings for the agency and advertiser to see before the final artwork is finished. This is seen as necessary as the costs of producing finished work and going live without any pre-testing can be critical, and expensive.

Storyboards are a way in which it is possible to inexpensively simulate a 'rough' version of the advertisement. Pen-and-ink line drawings, animatics or cartoons and photoboards are some of the more common approaches. Some storyboards will consist of as many as 20 sketches, depicting key scenes, camera and product shots, close ups, along with background scenery and essential props.

> Storyboards are a way in which it is possible to inexpensively simulate a 'rough' version of the advertisement.

Focus groups

When a small number (8–10) of target consumers are brought together and invited to discuss a particular topic a focus group is formed. By using in-depth interviewing skills a professional moderator can probe the thoughts and feelings held by the members of the group towards a product, media vehicles or advertising messages. One-way viewing rooms allow clients to observe the interaction without the focus group's behaviour being modified by external influences.

The advantage of focus groups is that they are relatively inexpensive to set up and run and they use members of the target audience. In this sense they are representative and allow true feelings and emotions to be uncovered in a way that other methods deny. They do not attempt to be quantitative and, in that sense, they lack objectivity. It is also suggested that the group dynamics may affect the responses in the 'artificial' environment. This means that there may be in-built bias to the responses and the interaction of the group members. Focus groups are very popular, but they should not be used on their own.

Consumer juries

A 'jury' of consumers, representative of the target market, is asked to judge which of a series of paste-ups and rough ideas would be their choice of a final advertisement. They are asked to rank in order of merit and provide reasons for their selections.

There are difficulties associated with ranking and prioritisation tests. First, the consumers, realising the reason for their participation, may appoint themselves as 'experts', so they lose

Exhibit 15.2 A focus group in operation
© Helen King/Corbis

> Emotional advertisements tend to receive higher scores than informational messages.

the objectivity that this process is intended to bring. Second, the halo effect can occur, whereby an advertisement is rated excellent overall simply because one or two elements are good and the respondent overlooks the weaknesses. Finally, emotional advertisements tend to receive higher scores than informational messages, even though the latter might do better in the marketplace.

Pre-testing finished advertisements

When an advertisement is finished it can be subjected to a number of other tests before being released.

Dummy vehicles

Many of the pre-testing methods occur in an artificial environment such as a theatre, laboratory or meeting room. One way of testing so that the reader's natural environment is used is to produce a dummy or pretend magazine that can be consumed at home, work or wherever participants normally read magazines. Dummy magazines contain regular editorial matter with test advertisements inserted next to control advertisements. These 'pretend' magazines are distributed to a random sample of households, which are asked to consume the magazine in their normal way. Readers are encouraged to observe the editorial and at a later date they are asked questions about both the editorial and the advertisements.

The main advantage of using dummy vehicles is that the setting is natural but, as with the focus group, the main disadvantage is that respondents are aware that they are part of a test and may respond unnaturally. Research also suggests that recall may not be the best measure for low-involvement decisions or where motivation occurs through the peripheral route of the elaboration likelihood model (ELM). If awareness is required at the point of sale, then recognition may be a more reliable indicator of effectiveness than recall.

Readability tests

Rudolph Flesch (1974) developed a formula to assess the ease with which print copy could be read. The test involves, among other things, determining the average number of syllables per 100 words of copy, the average length of sentence and the percentage of personal words and sentences. By accounting for the educational level of the target audience and by comparing results with established norms, the tests suggest that comprehension is best when sentences are short, words are concrete and familiar, and personal references are used frequently.

Projective techniques

Projective techniques are used to probe the subconscious and have close associations with Freudian thinking and the motivation school advocated by Dichter (1966) (see Chapter 5). Individuals or groups can be encouraged through projective techniques to express their inner thoughts and feelings about brands, products, services and organisations, among others. Four main projective techniques can be identified (see Table 15.2).

> Projective techniques are used to probe the subconscious.

Projective techniques have been used by many leading brands to understand how their brands are perceived, to test advertising and creative ideas and to segment their markets. For example, Guinness used projective techniques to understand how to position their brand and how advertising should be used to develop the ideal position.

Table 15.2 Projective techniques

Projective technique	Explanation
Association	Free word association tests require respondents to respond with the first word that comes to mind in response to a stimulus word. Often used in when naming brands.
Completion	Spontaneous sentence or story-telling completion are the most used methods. Responses can be graded as approval, neutral or disapproval enabling attitudes towards brands to be determined.
Transformation	These are also known as 'expressible' techniques and involve techniques such as psychodrawing. This requires respondents to express graphically their inner feelings about a brand or event (e.g. a shopping trip, holiday or purchase process).
Construction	This approach can involve role playing where respondents are asked to act out their feelings towards a purchase, a brand, event or organisation.

Source: Based on Robson (2002).

ViewPoint 15.1   Projective engagement with Sony Bravia

Projective testing techniques were used to measure the positioning success and the impact of the Sony Bravia 'Paint' ads. Sony wanted to use a creative that symbolised the technical colour development represented by the Bravia television.

Based on exploding colours around a council housing estate, Paint represented a radically different and unexpected creative, if only because there was no voice-over, no mention of attributes, features or benefits, in either copy or voice, and there were no visuals depicting the product or people consuming (watching) the television.

Paint was designed to communicate the point that the Bravia and SXRD range provide the 'colour that you'll see on these screens will be like no other' (www.Sony.com).

Part of the testing undertaken by TNS, using their AdEval™ methodology, included the use of people photosets. These are pictures of groups of personality types, used and validated internationally. Respondents were shown various different ads and asked which group of people they thought each would appeal to most. Many people have reported finding the Paint ads confusing and difficult to relate to. However, all the respondents were able to assign a personality type and the majority categorised the ads to people who were carefree, lively and bold. In terms of brand images the respondents reported vibrant colour, lively, outgoing, dynamic and cool/trendy as the key attributes (see Figure 15.1).

The claim made is that projective techniques can discover ways in which an ad engages with a consumer's emotions.
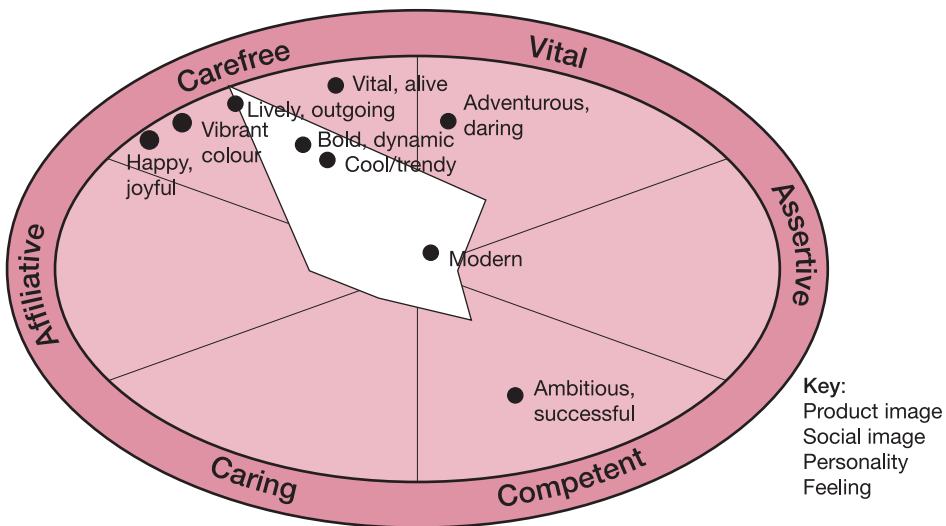
Source: Burden (2007).

Question

Which of the types of projective techniques set out in Table 15.2 do photosets best accompany?

Task

Try an informal word association test with your friends based around brands they enjoy.

The brand as portrayed by this ad has a lively, vital, dynamic, image; vibrant colour is clearly communicated

Figure 15.1   Results from projective technique for Sony Bravia television ad
Source: Burden (2007). Used with permission from WARC.

Theatre tests

As a way of testing finished broadcast advertisements, target consumers are invited to a theatre (laboratory or hall) to preview television programmes. Before the programme commences, details regarding the respondents' demographic and attitudinal details are recorded and they are asked to nominate their product preferences from a list. At the end of the viewing their evaluation of the programme is sought and they are also requested to complete their product preferences a second time.

> As a way of testing finished broadcast advertisements, target consumers are invited to a theatre (laboratory or hall) to preview television programmes.

There are a number of variations on this theme: one is to telephone the respondents a few days after the viewing to measure recall and another is to provide joysticks, push buttons and pressure pads to measure reactions throughout the viewing. The main outcome of this process is a measure of the degree to which product preferences change as a result of exposure to the controlled viewing. This change is referred to as the persuasion shift. This approach provides for a quantitative dimension to be added to the testing process, as the scores recorded by respondents can be used to measure the effectiveness of advertisements and provide benchmarks for future testing.

It is argued that this form of testing is too artificial and that the measure of persuasion shift is too simple and unrealistic. Furthermore, some believe that many respondents know what is happening and make changes because it is expected of them in their role of respondent. Those in favour of theatre testing state that the control is sound, that the value of established norms negates any 'role play' by respondents and that the actual sales data support the findings of the brand persuasion changes in the theatre.

A major evaluation of 400 individual advertising tests in the United States found, among many other things, that there is no clear relationship between measures of persuasion shift and eventual sales performance. This questions the use of an organisation's scarce resources and the viability of using these techniques (Lodish and Lubetkin, 1992).

This technique is used a great deal in the United States, but until recently has had limited use in Britain. However, Mazur (1993) reports that theatre testing is increasing in Britain. Agencies are concerned that the simplistic nature of recording scores as a means of testing advertisements ignores the complex imagery and emotional aspects of many messages. If likeability is an important aspect of eventual brand success then it is unlikely that the quantitative approach to pre-testing will contribute any worthwhile information.

The increasing use of, or at least interest in, theatre tests and the movement towards greater utilisation of quantitative techniques in pre-testing procedures runs concurrently with the increasing requirements of accountability, short-termism and periods of economic downturn. As no one method will ever be sufficient, a mix of qualitative and quantitative pre-test measures will, inevitably, always be required.

Physiological measures

A bank of physiological tests has been developed, partly as a response to advertisers' increasing interest in the emotional impact of advertising messages and partly because many other tests rely on the respondents' ability to interpret their reactions. Physiological tests are designed to measure the involuntary responses to stimuli that avoid the bias inherent in other tests. There are substantial costs involved with the use of these techniques, and the validity of the results is questionable. Consequently they are not used a great deal in practice, but, of them all, eye tracking is the most used and most reliable (see Table 15.3).

> Physiological tests are designed to measure the involuntary responses to stimuli that avoid the bias inherent in other tests.

Table 15.3   Physiological tests

Pupil dilation

Pupil dilation is associated with action and interest and is used to measure a respondent's reaction to a stimulus. If the pupil is constricted then interest levels are low and energy is conserved. The level of arousal is used to determine the degree of interest and preference in a particular advertisement or package design.

Eye tracking

This technique requires the use of eye movement cameras that fire an infrared beam to track the movement of the eye as it scans an advertisement. The sequence in which the advertisement is read can be determined and particular areas that do or do not attract attention can be located.

Galvanic skin response

This measures the resistance the skin offers to a small amount of current passed between two electrodes. Response to a stimulus will activate the sweat glands, which in turn will increase the resistance. Therefore the greater the level of tension induced by an advertisement the more effective it is as a form of communication.

Tachistoscopes

These measure the ability of an advertisement to attract attention. The speed at which an advertisement is flashed in front of a respondent is gradually slowed down until a point (about 1/100 second) is reached at which the respondent is able to identify components of the message. This can be used to identify those elements that respondents see first as a picture is exposed, and so facilitates the creation of impact-based messages.

Electroencephalographs

This involves the use of a scanner that monitors the electrical frequencies of the brain. Hemispheric lateralisation concerns the ability of the left-hand side of the brain to process rational, logical information and the right-hand side handles visual stimuli and responds more to emotional inputs.

Brain activation measures the level of alpha-wave activity, which indicates the degree to which the respondent is aroused by and interested in a stimulus. Therefore, the lower the level of alpha activity the greater the level of attention and cognitive processing. It would follow that, by measuring the alpha waves while a respondent is exposed to different advertisements, different levels of attention can be determined.

On the surface, pupil dilation has a number of attractions, but it is not used very much as research has shown little evidence of success. The costs are high and the low number of respondents that can be processed limits the overall effectiveness. Eye tracking can be a useful means of reviewing and amending the layout of an advertisement. Galvanic skin response is flawed because the range of reactions and emotions, the degree of learning and recall, and aspects of preference and motivation are all ignored. When these deficiencies are combined with the high costs and low numbers of respondents that can be processed, it is not surprising that this method of pre-testing has little value. The hemispheric lateralisation theory has been rejected by many researchers. Although the right side of the brain is best for recognition, and the left better for recall, only Vaughn (1980) has developed this approach in terms of advertising theory (Chapter 16).

Although now superseded, his grid was regarded as an important breakthrough in our understanding of how advertising works. However, while the grid has been used extensively, there is little evidence of any commercial application of electro-encephalographs. Advertisements should be designed to appeal to each hemisphere, but recent research now appears to reject this once-popular notion.

Post-testing

Testing advertisements that have been released is generally more time-consuming and involves greater expense than pre-testing. However, the big advantage with post-testing is that advertisements are evaluated in their proper environment, or at least the environment in which they are intended to be successful.

There are a number of methods used to evaluate the effectiveness of such advertisements, and of these inquiry, recall, recognition and sales-based tests predominate.

Inquiry tests

These tests are designed to measure the number of inquiries or direct responses stimulated by advertisements. Inquiries can take the form of returned coupons and response cards, requests for further literature or actual orders. They were originally used to test print messages, but some television advertisements now carry 0800 (free) telephone numbers. An increase in the use of direct-response media will lead to an increase in the sales and leads generated by inquiry-stimulating messages, so this type of testing will become more prevalent.

Inquiry tests can be used to test single advertisements or a campaign in which responses are accumulated. Using a split run, an advertiser can use two different advertisements and run them in the same print vehicle. This allows measurement of the attention-getting properties of alternative messages. If identical messages are run in different media then the effect of the media vehicles can be tested.

> Inquiry tests can be used to test single advertisements or a campaign.

Care needs to be given to the interpretation of inquiry-based tests, as they may be misleading. An advertisement may not be effective simply because of the responses received. For example, people may respond because they have a strong need for the offering rather than the response being a reflection of the qualities of the advertisement. Likewise, other people may not respond despite the strong qualities of the advertisement, simply because they lack time, resources or need at that particular moment.

Recall tests

Recall tests are designed to assess the impression that particular advertisements have made on the memory of the target audience. Interviewers, therefore, do not use a copy of the advertisement as a stimulus, as the tests are intended to measure impressions and perception, not behaviour, opinions, attitudes or the advertising effect.

Normally, recall tests require the cooperation of several hundred respondents, all of whom were exposed to the advertisement. They are interviewed the day after an advertisement is screened, hence the reference to day-after-recall (DAR) tests. Once qualified by the interviewer, respondents are first asked if they remember a commercial for, say, air travel. If the respondent replies 'Yes, Virgin', then this is recorded as unaided recall and is regarded as a strong measure of memory. If the respondent says 'No', the interviewer might ask the question 'Did you see an advertisement for British Airways?' A positive answer to this prompt is recorded as aided recall.

These answers are then followed by questions such as, 'What did the advertisement say about British Airways?', 'What did the commercial look like?' and 'What did it remind you of?' All the answers provided to this third group of questions are written down word for word and recorded as verbatim responses.

The reliability of recall scores is generally high. This means that each time the advertisement is tested, the same score is generated. Validity refers to the relationship or correlation between recall and the sales that ultimately result from an audience exposed to a particular advertisement. The validity of recall tests is generally regarded by researchers as low (Gordon, 1992).

> The reliability of recall scores is generally high.

Recall tests have a number of other difficulties associated with them. First, they can be expensive, as a lot of resources can be consumed by looking for and qualifying respondents. Second, not only is interviewing time expensive, but the score may be rejected if, on examination of the verbatim responses, it appears that the respondent was guessing.

It has been suggested by Zielske (1982) that thinking/rational messages appear to be easier to recall than emotional/feeling ones. Therefore, it seems reasonable to assume that recall scores for emotional/feeling advertisements may be lower. It is possible that programme content may influence the memory and lead to different recall scores for the same offering. The use of a preselected group of respondents may reduce the costs associated with finding a qualified group, but they may increase their attention towards the commercials in the knowledge that they will be tested the following day. This will inevitably lead to higher levels of recall than actually exist.

On-the-air tests are a derivative of recall and theatre tests. By using advertisements that are run live in a test area, it is possible to measure the impact of these test advertisements with DAR. As recall tests reflect the degree of attention and interest in the advertisement, this is a way of controlling and predicting the outcome of a campaign when it is rolled out nationally.

Recall tests are used a great deal, even though their validity is low and their costs are high. Wells *et al.* (1992) argue that this is because recall scores provide an acceptable means by which decisions to invest heavily in advertising programmes can be made. Agencies accumulate vast amounts of recall data that can be used as benchmarks to judge whether an advertisement generated a score that was better or less than the average for the product class or brand. Having said that, and despite their popularity, they are adjudged to be poor predictors of sales (Lodish and Lubetkin, 1992).

$\mathcal{V}iew\mathcal{P}oint$ 15.2   Total Recall

The techniques available to measure the ability of most media to deliver on core attributes such as reach and frequency is well rehearsed and understood. However, what is more problematic is measuring whether any particular variation or format within a type of media can impact on a viewer's/reader's attention or as some would have it, engagement with a medium and message.

Research has shown that people standing on London Underground stations and travelling in tube trains accumulate considerable more dwell time with respect to the ads they are exposed to. However, it was not known how dwell time might affect recall or the extent to which dwell time might impact on brand perceptions.

A research project, call Total Recall, was created. This exercise sought to measure the extent to which recall improves with length of exposure and whether positive perceptions (brand empathy) improves.

Using various poster ads, some complex with extensive copy and some using simple visual branding techniques, respondents were exposed to these ads and to the control ads within a managed testing environment. The results were quite clear. Respondents who were exposed to the ads for longer periods of time were more likely to remember what they had seen and were more likely to develop more positive feelings towards the brands.

Source: Cox (2007).

Question

Are the results of this research exercise surprising or could they have been confidently anticipated?

Task

Make a list of the ways in which people might be exposed to billboard and magazine ads more frequently?

Recognition tests

Recall tests are based upon the memory and the ability of respondents to reprocess information about an advertisement. A different way of determining advertising effectiveness is to ask respondents if they recognise an advertisement. This is the most common of the post-testing procedures for print advertisements. One of the main methods used to measure the readership of magazines is based on the frequency-of-reading and generally there are three main approaches:

> One of the main methods used to measure the readership of magazines is based on the frequency-of-reading.

- *recency:* reading any issue during the last publishing interval (e.g. within the last seven days for a weekly magazine);
- *specific issue:* reading of a specific issue of a publication;
- *frequency-of-reading*: how many issues a reader has read in a stated period (such as a month in respect of a weekly magazine).

Worldwide, the recency approach is the most widely used method in national readership surveys (www.roymorgan.com). Of the many services available, perhaps the Starch Readership Report is the best known. These recognition tests are normally conducted in the homes of approximately 200 respondents. Having agreed that the respondent has previously seen a copy of the magazine, it is opened at a predetermined page and the respondent is asked, for each advertisement, 'Did you see or read any part of the advertisement?' If the answer is yes the respondent is asked to indicate exactly which parts of the copy or layout were seen or read.

Four principal readership scores are reported: noted, seen-associated, read most and signature (see Table 15.4).

The reliability of recognition tests is very high, higher than recall scores. Costs are lower, mainly because the questioning procedure is simpler and quicker. It is also possible to deconstruct an advertisement into its component parts and assess their individual effects on the reader. As with all interviewer-based research, bias is inevitable. Bias can also be introduced by the respondent or the research organisation through the instructions given or through fatigue of the interviewer.

The validity of recognition test scores is said to be high, especially after a number of insertions. However, there can be a problem of false claiming, where readers claim to have seen an advertisement but, in fact, have not. This, it is suggested, is because when readers confirm they have seen an advertisement the underlying message is that they approve of and like that sort of advertisement. If they say that they have not seen an advertisement, the underlying message is that they do not usually look at that sort of advertisement. Wells *et al.* (1992) report on the view expressed by Krugman (1988), that readers are effectively voting on whether an advertisement is worth spending a moment of their time to look at. It might be that readers' memories are a

> The validity of recognition test scores is said to be high.

Table 15.4 Principal readership scores

Readership scores	Explanation
Noted	The percentage of readers who remember seeing the advertisement.
Seen-associated	The percentage of readers who recall seeing or reading any part of the advertisement identifying the offering.
Read most	The percentage of readers who report reading at least 50 per cent of the advertisement.
Signature	The percentage of readers who remember seeing the brand name or logo.

reliable indicator of what the reader finds attractive in an advertisement and this could be a surrogate indicator for a level of likeability. This proposition has yet to be fully investigated, but it may be that the popularity of the recognition test is based on the validity rating and the approval that high scores give to advertisers.

Sales tests

If the effectiveness of advertisements could be measured by the level of sales that occurs during and after a campaign, then the usefulness of measuring sales as a testing procedure would not be in doubt. However, the practical difficulties associated with market tests are so large that these tests have little purpose. Counting the number of direct response returns and the number of enquiries received are the only sales-based tests that have any validity.

Practitioners have been reluctant to use market-based tests because they are not only expensive to conduct but they are also historical by definition. Sales occur partly as a consequence of past actions, including past communication strategies, and the costs (production, agency and media) have already been sunk. There may be occasions where it makes little political and career sense to investigate an event unless it has been a success, or at the very least reached minimal acceptable expectations.

$ViewPoint$ 15.3   Ford measure their media

Measuring the return on their investment in marketing communications is really important to Ford. In addition to specific sponsorships Ford use a wide range of media, such as press, television, radio and online. It is critical that they understand what works and what does not work for them.

When Ford buy media space it is important that they achieve leverage. For example, when they launched the S-Max they took space in *The Observer* on- and offline as the readership profile matched that of the car.

Ford measure two key elements. The first concerns performance outputs or market measures such as market share at the retail level, profitability and ROI. They also measure the ROI on their direct mail and data-rich online activities. The second element concerns softer measures such as brand likeability, directly associated with some of their more emotionally oriented marketing goals.

However, the interesting aspect of Ford's media measurement activities concerns not what they measure but who actually does the measurement. All media buying and associated purchases are undertaken by MindShare but Ford then use the consulting company Accenture to measure the efficiency with which Mindshare buy media on Ford's behalf.

Source: Ovenden (2007).

Question

Why do Ford use Accenture to measure their own suppliers media buying effectiveness?

Task

Visit the Accenture web site www.accenture.com. How else do they contribute to the world of marketing communications?

For these reasons and others, advertisers have used test markets to gauge the impact their campaigns have on representative samples of the national market.

Simulated market tests

By using control groups of matched consumers in particular geographic areas, the use of simulated test markets permits the effect of advertising on sales to be observed under controlled

market conditions. These conditions are more realistic than those conducted within a theatre setting and are more representative of the national market than the limited in-house tests. This market representation is thought by some to provide an adequate measure of advertising effect. Other commentators, as discussed before, believe that unless advertising is the dominant element in the marketing mix, there are usually too many other factors that can affect sales. It is therefore unfair and unrealistic to place the sole responsibility for sales with advertising.

Single-source data

With the development and advances in technology it is now possible to correlate consumer purchases with the advertisements they have been exposed to. This is known as single-source data and involves the controlled transmission of advertisements to particular households whose every purchase is monitored through a scanner at supermarket checkouts. In other words, all the research data are derived from the same households.

The advent of cable television has facilitated this process. Consumers along one side of a street receive one set of control advertisements, while the others on the other side receive test advertisements. Single-source data provide exceptionally dependable results, but the technique is expensive, is inappropriate for testing single advertisements and tends to focus on the short-term effect, failing, for example, to cope with the concept of adstock.

In the United Kingdom facilities such as Adlab, then ScatScan and Homescan have helped advertisers assess their advertising effectiveness in terms of copy testing, weight testing and even the use of mixed media. The use of split regions can be very important, allowing comparisons to made of different strategies.

Other tests

There is a range of other measures that have been developed in an attempt to understand the effect of advertisements. Among these are tracking studies and financial analyses.

Tracking studies

A tracking study involves interviewing a large number of people on a regular basis, weekly or monthly, with the purpose of collecting data about buyers' perceptions of marketing communication messages, not just advertisements and how these messages might be affecting buyers' perceptions of the brand. By measuring and evaluating the impact of a campaign when it is running, adjustments can be made quickly. The most common elements that are monitored, or tracked, are the awareness levels of an advertisement and the brand, image ratings of the brand and the focus organisation, and attributes and preferences.

Tracking studies can be undertaken on a periodic or continuous basis. The latter is more expensive, but the information generated is more complete and absorbs the effect of competitor's actions, even if the effects are difficult to disaggregate. Sherwood *et al.* (1989) report that, in a general sense, continuous tracking appears more appropriate for new products and periodic tracking more appropriate for established products.

> Tracking studies can be undertaken on a periodic or continuous basis.

A further form of tracking study involves monitoring the stock held by retailers. Counts are usually undertaken each month, on a pre- and post-exposure basis. This method of measuring sales is used frequently. Audited sales data, market share figures and return on investment provide other measures of advertising effectiveness.

Tracking studies are also used to measure the impact and effectiveness of online activities. These may be applied to banner ads, email campaigns and paid-for search engine placements and have for a long time been geared to measuring site visitors, clicks through or pages visited. Increasingly these studies are attending to the volume and value of traffic with regard to the

behaviour undertaken by site visitors. Behaviour, or the more common term, call-to-action, can be considered in terms of the engagement through exchanges or transactions, the number of site or subscription registrations, the volume of downloads requested or the number of offline triggers such as 'call me buttons' that are activated.

Financial analysis

The vast amount of resources that are directed at planned communications, and in particular advertising, requires that the organisation reviews, on a periodic basis, the amount and the manner in which its financial resources have been used. For some organisations the media spend alone constitutes one of the major items of expenditure. For example, many grocery products incur ingredient, packaging and distribution plus media as the primary costing elements to be managed.

Variance analysis enables a continuous picture of the spend to be developed and acts as an early warning system should unexpected levels of expenditure be incurred. In addition to this and other standard financial controls, the size of the discount obtained from media buying is becoming an important and vital part of the evaluation process.

Increasing levels of accountability and rapidly rising media costs have contributed to the development of centralised media buying. Under this arrangement, the promotion of an organisation's entire portfolio of brands, across all divisions, is contracted to a single media-buying organisation. Part of the reasoning is that the larger the account the greater the buying power an agency has, and this in turn should lead to greater discounts and value of advertising spend.

The point is that advertising economies of scale can be obtained by those organisations that spend a large amount of their resources on the media. To accommodate this, centralised buying has developed, which in turn creates higher entry and exit barriers, not only to and from the market but also from individual agencies.

Likeability

A major study by the American Research Foundation investigated a range of different pretesting methods with the objective of determining which were best at predicting sales success. The unexpected outcome was that, of all the measures and tests, the most powerful predictor was likeability: 'how much I liked the advertisement'.

From a research perspective, much work has been undertaken to clarify the term 'likeability', but it certainly cannot be measured in terms of a simple Likert scale of 'I liked the advertisement a lot', 'I liked the advertisement a little', etc. The term has a much deeper meaning and is concerned with the following issues (Gordon, 1992):

- personally meaningful, relevant, informative, true to life, believable, convincing;
- relevant, credible, clear product advantages, product usefulness, importance to 'me';
- stimulates interest or curiosity about the brand; creates warm feelings through enjoyment of the advertisement.

The implication of these results is that post-testing should include a strong measure of how well an advertisement was liked at its deepest level of meaning.

> There are four main elements associated with likeability. These are entertainment, relevance, clearness (or clarity) and pleasantness.

Research by Smit *et al.* (2006) determined that there are four main elements associated with likeability. These are entertainment, relevance, clearness (or clarity) and pleasantness. Of these they found relevance to be the most important for changing viewer's opinions and entertainment for explaining how people process ads.

There are two main approaches to measuring likeability. One seeks to isolate what it is that viewers think and feel after seeing particular ads, i.e. how they feel. The other measures attitudes toward the ad itself. Essentially likeability is concerned with the affective element of the attitude construct. Indeed, some

researchers argue that likeability is a suitable response to the cognitive processing school of thought where individuals are considered to be rational problem-solvers.

Cognitive response analysis is an attempt to understand the internal dynamics of how an individual selects and processes messages, of how counter-arguing and message bolstering, for example, might be used to retain or reject an advertisement (see Chapter 17). Back in 1993, Biel reported a growing body of research evidence linking behaviour, attitude change and cognitive processing. He went on to predict, correctly, that this approach, unlike many of the others, was not restricted to FMCG markets and would be deployed across service markets, durables and retailers.

One of the important points to be made from this understanding of likeability is the linkage with the concept of 'significant value' considered in Chapter 8. The degree to which advertising works is a function of the level of engagement a message creates. This engagement is mediated by the context in which messages are sent, received and personally managed. The main factors are that the product in question should be new or substantially different, interesting, stimulating and personally significant. For advertising to be successful, it must be effective, and to be effective it should be of personally significant value to members of the target audience (those in the market to buy a product from the category in the near future).

The future use of technology will help the measurement and evaluation of advertising. The technology is now in place to meter what people are watching, by appending meters not to sets, but to people. Strapped-on mobile people meters can pick up signals indicating which poster site, television or radio programme is being walked past, seen or heard respectively.

Sales promotion

The measurement and evaluation of sales promotions are similar in principle to those conducted for advertising. The notion that some piloting should occur prior to launch in order that any wrinkles can be ironed out still holds strong, as does the need to balance qualitative with quantitative data. However, advertising seeks to influence awareness and image over the long term, whereas sales promotions seek to influence behaviour over the short term. As discussed earlier in this chapter, the evaluation of advertising can be imprecise and is subject to great debate. In the same way, the evaluation of sales promotions is subject to debate, but the means by which they are measured is not as ambiguous or as difficult as advertising (Shultz, 1987).

The use of quantitative methods as a testing tool leads to directly measurable and comparable outcomes, in comparison with the more subjective qualitative evaluations. Notionally, the balance in testing advertising is to use a greater proportion of qualitative than quantitative methods. The balance with sales promotions is shifted the other way. This is because the object being measured lends itself more to these kinds of measurement. If the purpose of sales promotion is to influence purchasing behaviour, then a measure of sales performance is necessary in addition to the evaluation of individual promotions.

> If the purpose of sales promotion is to influence purchasing behaviour, then a measure of sales performance is necessary in addition to the evaluation of individual promotions.

The different types of sales promotion are discussed in Chapter 24, where it is identified that there are a number of different target audiences for sales promotions activities; these are resellers, consumers and the sales force.

Manufacturer to reseller

The main objectives are to stimulate the resellers to try new products and to encourage them to allocate increased shelf space for established products. If campaigns are devised to meet these objectives, then a pre- and post-test analysis of the amount of allocated shelf space and the number of new products taken into the reseller's portfolio needs to be completed. These

processes are called retail audits (such as those undertaken by Nielsen Marketing Research), and although the information about changes in distribution and stock levels is not usually available until after the promotion has finished, it does provide accurate information concerning the effects that the event had on these variables.

Resellers to consumers

By generating higher levels of store traffic and moving stock from the store shelves to the consumers, sales promotions in this context require two main forms of evaluation. The first requires measures of the image held of the retailer, and this needs the use of tracking studies. The second requires measures of stock turnover per product category or brand against a predetermined planned level of turnover.

Manufacturers to consumers

The objectives are to encourage new users to try a product or to increase the amount that current users consume. Targets can be set for the number of coupons to be redeemed, sales generated during and after a price deal, the volume of bonus packs sold, the speed and volume of premiums disposed of and other direct measures of activity. Consumer audits reveal changes in the penetration and usage patterns of consumers. Redemption levels give some indication of participation levels, but should not be considered as the sole method of evaluation, as there are many people who might be encouraged to purchase by the promotion, but who then fail to participate for a variety of reasons.

> Redemption levels give some indication of participation levels.

Manufacturers to sales forces

The objectives of these activities are to build performance, morale and allegiance to the manufacturers and their products. Apart from measuring sales performance, the effectiveness of these activities can be expensive and difficult to measure. Attitude studies of the sales force can indicate the degree to which a contest has been influential, but it is hard to isolate the effects from those of other variables acting on them.

Through systematic tracking of sales and market share, products in mature markets can be evaluated in terms of their responsiveness to sales promotions. This type of information must be treated carefully, as the impact of other environmental factors has not been determined. Redemption rates allow for quantitative analysis, that, through time, leads to the establishment of a database from which benchmarks for promotional measurement and achievement can be obtained.

Using technology to evaluate sales promotions

It was noted in a previous section on advertising that advances in IT have radically altered the way in which advertising and product purchases can be evaluated. The same applies to sales promotions. It is now possible to predict with a high level of accuracy the impact on sales of different combinations of in-store promotions and price deals (Nielsen, 1993). This permits greater understanding of the way in which different sales promotions work and when they are most effective. This has two main benefits: the first is to focus promotions on activities that are effective; the second is to help target the communication spend on periods of the year, month and week that consumers are most responsive.

Homescan is an electronic household panel offered by ACNielsen that tracks day-to-day shopping patterns. It measures the household penetration and the retail distribution of a product. ACNielsen uses the system to analyse trial and repeat use and it provides data on consumer buying behaviour across most types of channel. These range from warehouse clubs and

convenience stores to supermarkets, mass merchandisers, mail order and the Internet. It can measure the number of households that use the product once and it can then determine how many of these trialists adopt a product through repeat purchase activity. It follows that test promotions can be used in particular stores or geographic areas, and control promotions can be used to test impact and effectiveness. What might work in one area might be unsuccessful elsewhere.

Coupons need not only be distributed via products and media. Technology has been developed that allows coupons of competitive brands to be automatically dispensed at the checkout once a product has been scanned. This information, together with the demographics and psychographic details compiled for panel members, enables detailed profiles to be built up about the types, timing and value of sales promotions to which different consumers respond.

Sales promotions are a competitive tool that allows for swift reaction and placement. In that sense, they are not being used as part of an overall campaign, more as an ad hoc sales boost. This implies that the manageability of sales promotions is very high relative to the other elements of the promotions mix and that the opportunity to pre-test might not be as large in practice as is theoretically possible (Peattie and Peattie, 1993).

The evaluation of sales promotion is potentially fast, direct, precise and easily comprehended (Doyle and Saunders, 1985). However, evaluation is not necessarily that clear cut. The synergistic qualities of the promotion mix inevitably lead to cross-over effects where the impact of other communications influences responses to particular sales promotion events. Promotions may also bring about increased awareness in addition to the trial, use and switching activities. Peattie and Peattie suggest that not only might brand and product substitution result from promotions, but store loyalty patterns might also be affected.

Of all the tools in the promotions mix, sales promotions lend themselves more easily to evaluation rather than to testing. Testing is not realistically possible in the time frames in which some organisations operate, particularly those in the FMCG sector. Activities should be planned and research built into campaigns, but it is the availability of improved IT that will continue to improve and accelerate the quality of information that management has about its sales.

> Of all the tools in the promotions mix, sales promotions lend themselves more easily to evaluation rather than to testing.

Public relations

Each of the two main forms, corporate and marketing public relations, seeks to achieve different objectives and does so by employing different approaches and techniques. However, they are not mutually exclusive and the activities of one form of public relations impact upon the others; they are self-reinforcing.

Corporate public relations (CPR)

The objectives that are established at the beginning of a promotional campaign must form the basis of any evaluation and testing activity. However, much of the work of CPR is continuous, and therefore measurement should not be campaign-oriented or time-restricted but undertaken on a regular, ongoing basis. CPR is mainly responsible for the identity cues that are presented to the organisation's various stakeholders as part of a planned programme of communications. These cues signal the visibility and profile of the organisation and are used by stakeholders to shape the image that each has of the focus organisation.

CPR is, therefore, focused on communication activities, such as awareness, but there are others such as preference, interest and conviction. Evaluation should, in the first instance, measure levels of awareness of the organisation. Attention should then focus on the levels of interest, goodwill and attitudes held towards the organisation as a result of all the planned and unplanned cues used by the organisation.

Traditionally these levels were assumed to have been generated by public relations activities. The main method of measuring their contribution to the communication programme was to collect press cuttings and to record the number of mentions the organisation received in the electronic media. These were then collated in a cuttings book that would be presented to the client. This would be similar to an explorer presenting an electric toaster to a tribe of warriors hitherto undisturbed by other civilisations. It looks nice, but what do you do with it and is it of any real use? Despite this slightly cynical interpretation, the cuttings book does provide a rough and ready way of appreciating the level of opportunities to see created by public relations activities.

> The cuttings book provides a rough and ready way of appreciating the level of opportunities to see created by public relations activities.

The content of the cuttings book and the recorded media mentions can be converted into a different currency. The exchange rate used is the cost of the media that would have been incurred had this volume of communication or awareness been generated by advertising activity. For example, a 30-second news item about an organisation's contribution to a charity event may be exchanged for a 30-second advertisement at rate card cost. The temptation is clear, but the validity of the equation is not acceptable. By translating public relations into advertising currency, the client is expected not only to understand, but also to approve of the enhanced credibility that advertising possesses. It is not surprising that the widely held notion that public relations is free advertising has grown so substantially when practitioners use this approach.

A further refinement of the cuttings book is to analyse the material covered. The coverage may be positive or negative, approving or disapproving, so the quality of the cuttings needs to be reviewed in order that the client organisation can make an informed judgement about its next set of decisions. This survey of the material in the cuttings book is referred to as a content analysis. Traditionally, content analyses have had to be undertaken qualitatively and were therefore subject to poor interpretation and reviewer bias, however well they approached their task. Today, increasingly sophisticated software is being used to produce a wealth of quantitative data reflecting the key variables that clients want evaluated.

Hauss (1993) suggests that key variables could include the type of publication, the favourability of the article, the name of the journalist, the audiences being reached and the type of coverage. All these and others can be built into programmes. The results can then be cross-tabulated so that it is possible to see in which part of the country the most favourable comments are being generated or observe which opinion formers are positively or negatively disposed.

Corporate image

The approaches discussed so far are intended to evaluate specific media activity and comment on the focus organisation. Press releases are fed into the media and there is a response that is measured in terms of positive or negative, for or against. This quality of information, while useful, does not assist the management of the corporate identity. To do this requires an evaluation of the position that an organisation has in the eyes of key members of the performance network. In addition, the information is not specific enough to influence the strategic direction that an organisation has or the speed at which the organisation is changing. Indeed, most organisations now experience perpetual change; stability and continuity are terms related to an environment that is unlikely to be repeated.

The evaluation of the corporate image should be a regular exercise, supported by management. There are three main aspects. First, key stakeholders (including employees, as they are an important source of communications for external stakeholders), together with members of the performance network and customers, should be questioned regarding their perceptions of the important attributes of the focus organisation and the business they are in (Chapter 12). Second, how does the organisation perform against each of the

> The evaluation of the corporate image should be a regular exercise.

attributes? Third, how does the organisation perform relative to its main competitors across these attributes?

The results of these perceptions can be evaluated so that corrective action can be directed at particular parts of the organisation and adjustments made to the strategies pursued at business and functional levels. For example, in the computer retailing business, prompt home delivery is a very important attribute. If company A had a rating of 90 per cent on this attribute, but company B was believed to be so good that it was rated at 95 per cent, regardless of actual performance levels, then although A was doing a superb job it would have to improve its delivery service and inform its stakeholders that it was particularly good at this part of the business.

Recruitment

Recruitment for some organisations can be a problem. In some sectors, where skills are in short supply, the best staff gravitate towards those organisations that are perceived to be better employers and provide better rewards and opportunities. Part of the task of CPR is to provide the necessary communications so that a target pool of employees is aware of the benefits of working with the focus organisation and develops a desire to work there.

Measurement of this aspect of CPR can be seductive. It is tempting just to measure the attitudes of the pool of talent prior to a campaign and then to measure it again at the end. This fails to account for the uncontrollable elements in CPR, for example the actions of others in the market, but, even if this approach is simplistic and slightly erroneous, it does focus attention on an issue. A major chemical-processing company found that it was failing to attract the necessary number of talented undergraduates, partly because the organisation was perceived as unexciting, bureaucratic and lacking career opportunities. A coordinated marketing communications campaign was targeted at university students, partly at repositioning the organisation in such a way that the students would want to work for them when they finished their degrees. The results indicated that students' approval of the company as a future employer rose substantially in the period following the campaign.

Crisis management

During periods of high environmental turbulence and instability, organisations tend to centralise their decision-making processes and their communications (Quinn and Mintzberg, 1992). When a crisis occurs, communications with stakeholders should increase to keep them informed and aware of developments. In Chapter 19, it will be observed that crises normally follow a

> When a crisis occurs, communications with stakeholders should increase to keep them informed and aware of developments.

number of phases, during which different types of information must be communicated. When the crisis is over, the organisation enters a period of feedback and development for the organisation. 'What did we do?', 'How did it happen?', 'Why did we do that?' and 'What do we need to do in the future?' are typical questions that socially aware and mature organisations, which are concerned with quality and the needs of their stakeholders, should always ask themselves.

Pearson and Mitroff (1993) report that many organisations do not expose themselves to this learning process for fear of 'opening up old wounds'. Those organisations that do take action should communicate their actions to reassure all stakeholders that the organisation has done all it can to prevent a recurrence, or at least to minimise the impact should the origin of the crisis be outside the control of management. A further question that needs to be addressed concerns the way the organisation was perceived during the different crisis phases. Was the image consistent? Did it change, and if so why? Management may believe that it did an excellent job in crisis containment, but what really matters is what stakeholders think – it is their attitudes and opinions that matter above all else.

The objective of crisis management is to limit the effect that a crisis might have on an organisation and its stakeholders, assuming the crisis cannot be prevented. The social system in which an organisation operates means that the image held of the organisation may well change

as a result of the crisis event. The image does not necessarily become negative. On the contrary, it may be that the strategic credibility of the organisation could be considerably enhanced if the crisis were managed in an open and positive way. However, it is necessary that the image that stakeholders have of an organisation should be tracked on a regular basis. This means that the image and impact of the crisis can be monitored through each of the crisis phases. Sturges *et al.* (1991) argue that the objective of crisis management is to influence public opinion to the point that 'post-crisis opinions of any stakeholder group are at least positive, or more positive, or not more negative than before the crisis event'. This ties in with the need to monitor corporate image on a regular basis. The management process of scanning the environment for signals of change, along with change in the attitudes and the perception held by stakeholders towards the organisation, make up a joint process that public relations activities play a major role in executing.

Marketing public relations (MPR)

It was identified earlier that there is evidence of the increasing use of MPR. There are many reasons for this growth, but some of the more important ones quoted by organisations are rising media costs, audience fragmentation, changing consumer attitudes and increasing educational needs (Kitchen, 1993). By using public relations to support the marketing effort in a direct way, organisations are acknowledging that the third-party endorsement provided by MPR delivers a high level of credibility and cost effectiveness, which the other elements of the promotions mix fail to provide.

As Kitchen rightly argues, MPR cannot exist in a vacuum; it must be integrated with the other elements of the mix and provide complementarity. It is the use of MPR as a form of product support and as part of a planned communications mix that makes this a source of high-quality leads. However, evaluating the contribution of MPR is problematic.

Some practitioners believe that this can be overcome by coding press releases as a campaign, and with the use of particular software, leads can be tracked and costed. With the right software, the actual cost of a press release can be input and the number of leads that come back can be measured against sales on the database.

The software can not only estimate sales but also work out the number of leads required to make quota. The formula used is based on the rule that 45 per cent of leads turn into sales for someone in the market within the year. The organisation's own conversion rate can be used to adjust the 45 per cent and the quality of its lead conversion process can also be input. One of the benefits of this approach is that quantitative outcomes provide a measure of effectiveness, but not necessarily the effectiveness of the MPR campaign.

> Pre- and post-test measures of awareness, preference, comprehension and intentions are a better measure of quality and impact.

Pre- and post-test measures of awareness, preference, comprehension and intentions are a better measure of the quality and impact that an MPR campaign might have on a target audience. Measuring the conversion ratio of leads to sales is not the only measure, as it fails to isolate the other forces that impact on market performance.

MPR in business-to-business markets is directly targeted at members of the performance network. The objectives are many and include building awareness, reducing costs, satisfying educational needs and enhancing image through improving credibility. The overriding need, however, is to improve the relationship between members

> Measurement of the effectiveness of MPR should be undertaken by evaluating the degree to which members support, like, endorse or prefer an organisation.

of the network and to provide them with a reason to continue transactions with the focus organisation. The reasons are similar to those in the personal selling/buying formula (Chapter 22), namely to associate product adequacy when the appropriate problem is surfaced and to create pleasant feelings when the name of the product or the organisation is mentioned in the same context. MPR in this situation is being used as a competitive tool

to defend established positions. Measurement of the effectiveness of MPR, therefore, should be undertaken by evaluating the degree to which members support, like, endorse or prefer an organisation and the products it offers. This can be achieved through the use of tracking studies that plot attitudes and opinions, against which the timings of campaigns and MPR activities can be traced and evaluated.

Other measuring techniques – PR

Of all the tools available to practitioners, Goften (1999) reports the following as the most common approaches to measuring public relations:

- *set objectives* and agree the criteria in advance of a campaign;
- *press cuttings*, radio and television tapes, but this is a measure of volume and not quality of impact. A media equivalent value is then applied;
- *media evaluation* through commercial systems such as CAMMA, impact, precis. Under this approach, panels of readers judge whether a mention is positive or negative and whether the client's key message has been communicated. Computer programmes then cut through the data;
- *tracking studies* are expensive but important when changing a perception of a brand, etc.

Both CPR and MPR are difficult and elusive elements of the promotional mix to test, measure and evaluate. Practitioners use a variety of methods, but few provide the objectivity and validity that is necessary. For example, Comic Relief monitored the impact of media coverage on the organisation in the run-up to Red Nose Day. It was able to track which initiatives were failing to attract attention and which issues were attracting negative coverage. It evaluated coverage over six key areas: television initiatives, education, grants (Africa and the United Kingdom), special projects, public fundraising and corporate fundraising.

The variety of measurement devices is increasing, especially as technology advances.

ViewPoint 15.4 Watching over TeleTrax

TeleTrax is an electronic tagging system that can monitor broadcast use of its footage. Using an indelible code embedded within video tapes and through the use of approximately 100 listening posts across Europe, the company is alerted as soon as the tape is broadcast. Unfortunately it does not track the tone of the content. Monitoring the Internet for PR coverage is more taxing.

Net.Cut was set up originally to provide early warning of unfavourable corporate comment on the Internet. It can monitor comment in Internet publications, UK newsgroups and the world-wide web by searching the world-wide web at night and saving company mentions. The cuttings are then reviewed the following morning for key messages prior to warning the client as necessary.

Question

What advantages do you believe TeleTrax offers clients?

Task

Make a list of the types of information you would like TeleTrax to find.

If, at the end of the process, evaluation and testing lack objectivity, then the method should not be used. As a greater number of organisations are beginning to recognise the impact that public relations can provide and establish a more credible balance for the promotional mix, so there is a greater requirement for planning and evaluation to be built into the process from the beginning (Watson, 1992).

Sponsorship

Both academics (Armstrong, 1998) and practitioners (Ovenden, 2007) agree that the measurement of sponsorship activities is problematic, although the importance of doing so is recognised and accepted. The problem concerns the ability to separate the impact of the various elements of the promotional mix, which can be expensive and beyond the reach of smaller brands.

Many organisations attempt to measure the size of the media audience and then treat this as an indicator of effectiveness. This is misleading, as advertising and sponsorship are considered to work in different ways and cannot be measured in a similar way. Audiences consider events (a sports match, exhibition or television programme) as their primary focus, not which organisation is sponsoring the activity, unlike advertising, where the message either dominates the screen or a page of a magazine and viewers attend according to their perceptual filters. The focus of attention is different, as should be the means of evaluation.

Marshall and Cook (1992) found that sports sponsors preferred to use consumer surveys to examine customer (not audience) profiles, brand-related images, attitudes and purchasing activities. This was accomplished through the use of personal interviews and telephone and postal surveys. Because the level of funding in many of the smaller sponsorships is relatively low, few if any resources are allocated to evaluative practices.

Taylor Nelson Sofres provides a single-source data panel through which the viewing habits and purchase behaviour of a representative panel of consumers are monitored through a system called TVSpan 3000. Unlike similar competitive offerings, TVSpan 3000 uses data from 3,000 homes where television set meters are installed that monitor each household's live television viewing, minute by minute. These homes are also equipped with AGB Superpanel scanning equipment for recording details of purchases of FMCG products. One of its prime tasks is to enable clients to monitor purchase behaviour and test advertisements at either pre- or post-test stage. Further uses are to test new creative ideas, to test the effects of advertising with or without below-the-line support and, interestingly, to test the interaction between advertising and sponsorship (Thorncroft, 1996).
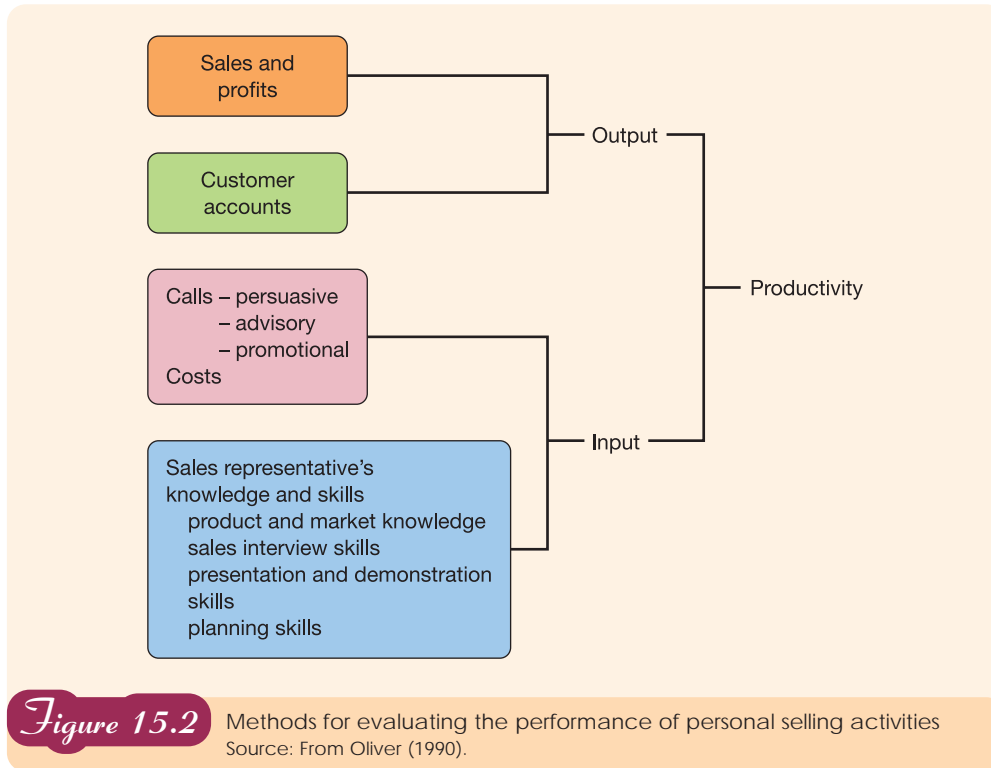
> The main way in which sponsorship activities should be measured is through the objectives set at the outset.

The main way in which sponsorship activities should be measured is through the objectives set at the outset. By measuring performance rigorously against clearly defined sales and communication-based measures it is more likely that a reasonable process and outcome to the sponsorship activity will be established.

Personal selling

In contrast to the other elements of the promotional mix, personal selling requires different methods of evaluation. Pre- and post-testing the performance of each salesperson is impractical and inappropriate. What is more pertinent for evaluation are the inputs and the effectiveness (measured as outcomes) of the personal selling process. Oliver (1990) suggests that performance can be seen as a factor of the effort and costs (inputs) that an organisation contributes. Outputs can be regarded as sales and profits resulting from exchanges with customers, while productivity can be deemed to be the ratio of inputs to outputs (see Figure 15.2).

This is a useful approach because it focuses attention on aspects of the promotional process that can be measured with the use of quantitative tools. This contrasts with the other tools, where qualitative measures generally predominate. In addition to this framework, it is necessary to measure the effectiveness of the sales force as a unit and the degree to which interdepartmental cooperation is achieved in synchronising the activities of the mix.

Figure 15.2 Methods for evaluating the performance of personal selling activities
Source: From Oliver (1990).

Evaluating the performance of a salesperson

The performance of a salesperson requires the use of both qualitative and quantitative methods. There are two main types of inputs to the sales process. The first of these consists of the activities undertaken and the costs incurred as a result. The second type of input concerns the knowledge and skills necessary to achieve the required outputs. These will be examined in turn.

Measuring and then evaluating the activities of each salesperson, the inputs, is an important and frequently used measuring stick. The number of planned and unplanned sales calls, the number of presentations, the frequency with which the showroom has been used and the mix of accounts visited, plus the expenses, cost of samples used and time associated with these activities, can be measured and evaluated against organisational standards and expectations. These simple quantitative measures provide for objectivity and measurement; what they do not do is provide an insight into why the input and the ultimate performance rating did or did not achieve the required standard.

Measuring and evaluating the knowledge component of the input dimension require greater subjectivity and reliance on qualitative measures. How well a sales-person uses their selling skills and presents themselves to customers is vitally important. In addition, the depth of knowledge that the subject has of the products, customers, territory and market will probably have a greater bearing on the performance outcome than the number of visits made. In other words, it is the quality of the sales call that is important, not the number of sales calls made. The measurement of these qualitative aspects relies on the individual judgement of the people responsible for the evaluation process (Churchill *et al.*, 1990).

Outputs are more easily measured than inputs. The most common technique used is that of the ratings attached to the volume or value of sales generated in a particular period in a designated area. Using a quota to measure achievement can be important for consistent tracking of performance and for motivational purposes. Volume analysis allows management to measure the effectiveness of the

> Outputs are more easily measured than inputs.

sales process, as comparisons can be drawn with last year's performance, with other salespersons (with similar territory potential) and with the potential in the territory.

Ratios provide a further insight into the overall performance and productivity of a salesperson. Expense ratios are a useful tool for understanding the way in which a salesperson is managing the territory. The cost/call ratio, for example, reveals the extent to which the subject is making calls and the costs of supporting the individual in the territory. Further detailed analyses are possible, for example travel expenses/call:

$$\text{sales expense ratio} = \frac{\text{expenses}}{\text{sales}}$$

$$\text{cost per call ratio} = \frac{\text{total}}{\text{number of calls}}$$

Servicing ratios reveal the extent to which a territory's business potential has been acquired, for example, what percentage of a territory's accounts has been won, how many prospects become customers, how many customers are lost and what level of sales is achieved on average per customer or per call:

$$\text{account penetration ratio} = \frac{\text{accounts sold into}}{\text{total number of available accounts}}$$

$$\text{average order size ratio} = \frac{\text{total sales value}}{\text{total number of orders}}$$

$$\text{new account ratio} = \frac{\text{number of accounts}}{\text{total number of accounts}}$$

The final group, activity ratios, determines the effort that is put into a territory. Calls/day, calls/account type and orders/call reveal the amount of planning and thought that is being put into an area:

$$\text{calls/accounts ratio} = \frac{\text{number of calls made}}{\text{total number of accounts}}$$

$$\text{orders/calls ratio} = \frac{\text{number of orders}}{\text{total number of calls}}$$

$$\text{calls/week ratio} = \frac{\text{number of calls}}{\text{number of weeks worked}}$$

In isolation these ratios provide some objectivity when attempting to measure the performance of a salesperson. Used in combination they become a more powerful tool, but only to the extent that they are an aid to decision-making. One major advantage of ratio analysis is the benchmarking effect. Comparisons become possible not only across the sales force but also across the industry, as norms become established through time.

While the traditional measure has been volume, emphasis has been increasingly placed on measures of profitability, an efficiency measure. The level of gross margin achieved by each salesperson and the contribution each makes to the overall profitability of the organisation are regarded by many organisations as more important than measures of volume. The approach requires the involvement of each salesperson not only in achieving the outcomes but also in the process of setting the appropriate performance targets in the first place. This requires different types of training and skills development, which in turn will affect the expectations held by each member of the sales force.

Evaluating the performance of a sales force

The methods looked at so far have been used to evaluate the performance of individual salespersons. An overall measure of the effectiveness of the larger unit, the sales force, is also

necessary. The following constitute the main areas of evaluation: the objectives set in the promotion mix, the level of interaction with the other elements of the promotion mix, activity measures and achievement against quota, the effectiveness of the sales channels used and the quality of the relationships established with customers.

The sales force, as a part of the promotion mix, has a responsibility to achieve the sales objectives set out in the promotion objectives. To do this the sales force needs the support of the other elements of the mix. Measuring this interdisciplinary factor is extremely difficult, but there is no doubt that each of the elements works more efficiently if they are coordinated with one another and the messages conveyed dovetail and reinforce each other.

> The sales force has a responsibility to achieve the sales objectives.

Many of the measures used to evaluate the performance of individual salespersons can also be aggregated and used to evaluate the performance of the sales force as a whole. The sales force will have an overall sales budget, usually by volume and value, against which actual performance can be measured. The sales force will also be expected to open an agreed number of accounts each period and the value of business as a proportion of the potential will be watched closely.

There is no doubt that the role of the sales force is changing. If the expectations of the sales force are being adapted to new environmental conditions, it is probable that alternative measures will be required to determine the progress that a sales force is making. For example, in business-to-business markets, the traditional approach of the sales force is to manage products and their allocation to selected customers. The sales force of the future is going to be responsible, to a much greater degree, for the management of customer relationships (Wilson, 1993) and the maintenance of relational transactions that will provide organisations with strategic advantage. The use of simple quantitative techniques to measure the performance of the sales force will decline, the use of qualitative techniques will become more prevalent and the techniques themselves will become sophisticated. Measures will be required to evaluate the quality of the relationships developed by the sales force rather than the quantity of outputs achieved in a particular period. The traditional emphasis on short-term quota achievement may well change to a focus on long-term customer alliances and an evaluation of the strength of the relationship held between partners.

One further area of evaluation that is necessary is that of the sales channels themselves. The increased use of multiple sales channels and the contribution that direct marketing will make to the sales force cannot be ignored. Measures are required of the effectiveness of the field sales force, the key account selling team and the array of direct marketing techniques. Constant monitoring of the market is required to judge whether the classification of an account should be changed, and whether different combinations of selling approaches should be introduced.

Finally, customers need to be involved in the sales channel decision process and in the evaluation of the field sales force. If customers are happy with a sales channel, then they are more likely to continue using it. It is vital that the views of customers are monitored regularly and that they contribute to the evaluation process.

The evaluation of the sales force and its individual members has long been oriented to quantitative measures of input and output productivity. These are useful, as they provide for comparison within the organisation and with the industry norms. However, in future, evaluation will move from a revenue to a profit perspective and much greater emphasis will be placed on the quality of the relationships that the sales force develops with their customers. The current imbalance between the use of quantitative and qualitative measures will shift to a position where qualitative measures become more important in evaluating the performance of the sales force.

> Evaluation will move from a revenue to a profit perspective.

Measuring the fulfilment of brand promises

Brands make promises and communicate them in one of two main ways. One is to make loud claims about the brand's attributes and the benefits these deliver to customers. This approach tends to rely on advertising and the strength of the brand to deliver the promise. The alternative is not to shout, but to whisper, and then surprise customers by exceeding their expectations when they experience the brand. This is an under promise/over deliver strategy, one which reduces risk and places a far greater emphasis on word-of-mouth communication, and brand advocacy. This in turn can reduce an organisation's investment in advertising and lead to a redirection of communication effort and resources in order to improve the customer experience.

It follows therefore that there are measurable gaps between the image and perceptions customers have of brands and their actual experiences. Where expectations are exceeded the promise gap is said to be positive. Where customers feel disappointed through experience of a brand, a negative promise gap can be identified. These gaps are reflected in the financial performance of brands.

The Promise Index reported by Simms (2007) found that although 66 per cent of the brands surveyed had positive promise gaps only 15 per cent had gaps that impacted significantly on business performance. Other research by Weber Shandwick found that the main factor for creating brand advocacy was the ability to 'surprise and delight customers'. This survey of 4,000 European consumers, reported by Simms, found that brand advocacy is five times more likely to prompt purchase than advertising.

A related metric, the Net Promoter Score (NPS), seeks to identify how likely an individual is to recommend a brand. Again, a key outcome is that brand growth is driven principally by surprising and delighting customers.

On the basis that brand advocacy is of major importance, two key marketing communication issues emerge. The first concerns how the marketing communications mix should be reformulated in order to encourage brand advocacy. It appears that advertising and mass media have an important role to play in engaging audiences to create awareness and interest. However, more emphasis needs to be given to the other tools and media in order to enhance each customer's experience of a brand beyond their expectations.

The second issue concerns identifying and communicating with passive rather than active advocates. Encouraging customers to talk about a brand means developing content that gives passive advocates a reason to talk about a brand. This means that the message component of the mix needs to be designed away from product attributes and towards stories and memorable events that can be passed on through all customer contact points. This in turn points to a greater use of public relations, viral and the use of user-generated-content, networks and communities and the use of staff in creating brand experiences.

ViewPoint 15.5   Measuring gorillas

For Cadbury the news had not been very good in 2006 and the first part of 2007. A salmonella scare in 2006 had resulted in more than 1 million bars of chocolate (including Dairy Milk) being removed from shelves accompanied by considerable negative media comment. In Februrary 2007, Cadbury spent £10 million launching the US chewing-gum Trident but then became enmeshed in controversy after the Advertising Standards Authority reprimanded the company for an ad featuring an Afro-Caribbean poet which was perceived to be racially offensive. This was followed by the news that Cadbury Schweppes was having to shed staff in a £300 million cost-cutting exercise.

In the autumn however it launched an ad that propelled its reputation in the opposite direction. The ad featured a man in a gorilla suit playing the drums to the Phil Collins hit 'In the Air Tonight'.

The ad caught the public's imagination, if only because there is no reason for a gorilla to play the drums, there is no connection between Cadbury's and a gorilla (at the time) and the ad says nothing about Dairy Milk apart from a shot of the brand name at the end. The ad was relatively inexpensive to produce and was released through a spoof real film production company, A Glass and a Half Full Productions.

The ad featured in a pre-ad teaser campaign in television listings that resembled a film. The Glass and a Half Full Productions web site helped to sustain dialogue with fans while 90-second spots during the Rugby World Cup and Big Brother Finals delivered the ad to huge audiences.

The ad was not part of an integrated campaign and the support might have been stronger. However, it was a one-off masterpiece of creativity that resonated with the nation. Sales rose 7 per cent by the end of October 2007 in value terms, and weekly sales were up 9 per cent year on year during the period 'gorilla' was on air. The ad generated the highest recognition scores ever recorded by Hall & Partners.

Sources: Various including Campaign (2007).

Question

How should the 'Gorilla' ad be evaluated, and how would you measure its success?

Task

Gorillas feature in other ways for some other brands. Find two other campaigns that feature gorillas.

Exhibit 15.3 The Cadbury gorilla in action
Image courtesy of The Advertising Archives.

Online communications

Online research has grown as the Internet population has soared and the measures used have developed through trial and experience. However, the notion that measurement of online communications is easy simply because all that is necessary is 'counting clicks' is misleading. Indeed, when speaking about marketing communications, Roisin Donnelly, UK & Ireland Corporate Marketing Director and Head of Marketing for Procter & Gamble, states that measurement is their biggest difficulty. It is not an issue to measure the overall impact of a campaign but when measuring the return generated by integrated campaigns, the contribution that word-of-mouth, blogging, online and public relations make, for example, is extremely difficult to isolate (Lannon, 2007).

Banner ads

> Traditional measurement techniques of reach, frequency and target audience impressions are not capable of being readily transferred to the Internet.

Not surprisingly there is disagreement about whether it is possible to measure effectively online advertising. Dreze and Zurfryden (1998) rightly point out that as a viable advertising media, Internet advertising must be subject to suitable measurement standards to gauge the effectiveness of the medium. Web servers can indicate how many pages have been requested, the time spent on each page and even the type of computers that were used to request the page. However, this type of information is largely superficial and fails to provide insight into the user, their motivation to visit the site or the behavioural or attitudinal outcomes as a result of the interaction. Traditional measurement techniques of reach, frequency and target audience impressions are not capable of being readily transferred to the Internet.

Others argue that it is possible to measure online tools. For example, Briggs and Hollis (1997) point out that one of the more common measures used is the click-through rate. They indicate, however, that this normally only measures behaviour, whereas what is needed is an indicator of the user's attitudes. They claim to have developed a technique to measure attitudes (online) and show that banner advertising can be one of the most effective forms of advertising and brand development.

Web site effectiveness

Johnston cited by Gray (2000) reports that ACNielsen offers a 9,000 strong panel in the United Kingdom. The panel consists of Internet users who have special software loaded on their PCs that records every web page they visit. The strengths and weaknesses of online qualitative and quantitative research are shown in Tables 15.5 and 15.6.

Good marketing management practice suggests that evaluation of any management activity should always include a consideration of the degree to which the objectives have been satisfied.

However, the reasons organisations have for setting up a web site are many and varied: these might be to establish a web presence, to move to new methods of commercial activity, to enter new markets, to adhere to parent company demands or to supplement current distribution channels. Consequently, it is not practicable to set up a definitive checklist to use as a measure of web site effectiveness, although certain principles need to be followed (see Table 15.7).

One of the basic approaches is to develop profiles of web site visitors built up by presenting every tenth visitor with a questionnaire. The next stage will be to provide media planners with these data to optimise banner ad placement. Based on the work of Berthon *et al.* (1996), Table 15.6 suggests the criteria that might be used to test a site's effectiveness, but different criteria will have a different impact depending upon each organisation's situation.

Table 15.5   Online quantitative research

Strengths	Weaknesses
Relatively inexpensive	Respondent universe
Fast turnaround	Sampling issues: narrow target audience and difficult to identify
Automated data collection	
Can show graphics and video	Often self-completion, hence subject to self-selection
No interviewer bias	
Quality of data	Technical problems
Seamless international coordination	

Table 15.6   Online qualitative research

Strengths	Weaknesses
Slightly faster and cheaper than traditional focus group	Loss of non-verbal communications
	Less useful for emotional issues
Avoids the dominance of loud personalities	Online moderation requires new skills pattern
More client control	Slow keyboard skills can hamper some respondents
Can show concepts and/or web sites	Sampling issues: difficult to identify a narrow target audience
Allows for international coordination and permits mixed nationalities	

Table 15.7   Criteria to assess web site effectiveness

Visitor type	Cognitive state	Management action
All surfers	Level of awareness that a site exists: aware or not aware	Provide offline and online information and directions
Those aware	Level of interest in the site: interested or not interested	Create interest and curiosity
Those interested	Known route to the site: determined or accidental	Enable greater opportunities for site hit
Determined visitors	Was the visit completed successfully? Transaction or no transaction	Encourage bookmarking and post-purchase communication to permit legitimate dialogue
Those who transacted	Will these visitors return to the site? Retained or not retained	Maintain and enhance top-of-mind site recall

Source: Adapted from Berthon *et al.* (1996). Used with kind permission from WARC.

Dreze and Zurfryden (1998) were apprehensive of the difficulties associated with measuring the number of unique site visitors, mainly because of various technology-related factors and the difficulties of isolating who is a unique visitor.

As mentioned earlier in this chapter, tracking studies are used to measure online brand values (site visits) and various forms of calls-to-action.

Summary

In order to help consolidate your understanding of the methods and approaches used to evaluate marketing communications, here are the key points summarised against each of the learning objectives:

1. Discuss the role of evaluation as part of marketing communications.

The evaluation of a marketing communications plan, once implemented, is an essential part of the total system. The evaluation provides a potentially rich source of material for the next campaign and the ongoing communications that all organisations operate, either intentionally or not.

The evaluation of planned marketing communications consists of two distinct elements. The first element is concerned with the development and testing of individual messages. For example, a particular sales promotion (such as a sample pack) has individual characteristics that may or may not meet the objectives of a sales promotion event.

The second element concerns the overall impact and effect that a campaign has on a target audience once a communications plan has been released. This post-test factor is critical, as it will either confirm or reject management's judgement about the viability of its communications strategy. The way in which the individual components of the communications mix work together needs to be understood so that strengths can be capitalised on and developed and weaknesses negated.

2. Explore the value and methods of pretesting and post-testing advertisements.

Pretesting is the practice of showing unfinished commercials to selected groups of the target audience with a view to refining the commercial to improve effectiveness, it is still subject to debate about its effectiveness. Many practitioners feel that pretesting limits creativity. The most common methods used to pretest advertisements are concept testing, focus groups, consumer juries, dummy vehicles, projective assessments, readability, theatre and physiological tests.

Post-testing is the practice of evaluating ads that have been released. The main advantage with post-testing is that advertisements are evaluated in their proper environment. There are a number of methods used to evaluate the effectiveness of such advertisements, and of these inquiry, recall, recognition and sales-based tests predominate.

3. Explain the main ideas behind different physiological measures of evaluation.

A bank of physiological tests has been developed, partly as a response to advertisers' increasing interest in the emotional impact of advertising messages and partly because many other tests rely on the respondents' ability to interpret their reactions. Physiological tests have been designed to measure the involuntary responses to stimuli and so avoid the bias inherent in the other tests. There are substantial costs involved with the use of these techniques, and the validity of the results is questionable. Consequently they are not used a great deal in practice, but, of them all, eye tracking is the most used and most reliable.

4. Provide an insight into the way in which each of the tools of the communication mix can be evaluated.

Each of the tools of the mix require different forms of evaluation or testing, simply because of the way they work. Sales promotion and direct marketing provide relatively easy forms of measurement and evaluation based on quantitative methods. Public relations however, is far more subjective and measurement is much more problematic. The effectiveness of personal selling can be measured at an individual sales person level or in aggregate form, the sales team level.

5. Consider other ways in which the effectiveness of marketing communications can be evaluated.

Tracking studies involve interviewing a large number of people on a regular basis, weekly or monthly, with the purpose of collecting data about buyers' perceptions of marketing communication messages. Perceptions, attitudes and meanings attributed to campaigns can be tracked and adjustments made to campaigns as necessary.

Considering the vast amount of resources that are directed at planned communications, and in particular advertising, it is important to review, on a periodic basis, the amount and the manner in which its financial resources have been used. For some organisations the media spend alone constitutes one of the major items of expenditure.

Likeability refers to the deep seated set of meanings and associations an individual makes with particular ads and its measurement has proven to be a good indicator of an ad's performance.

6. Measure the fulfilment of brand promises.

There are measurable gaps between the image and perceptions customers have of brands and their actual experiences. Where expectations are exceeded the promise gap is said to be positive. Where customers feel disappointed through experience of a brand, a negative promise gap can be identified. These gaps are reflected in the financial performance of brands.

The Promise Index and the Net Promoter Score (NPS) are two approaches to measuring the success of delivering brand promises.

7. Consider some of the issues associated with evaluating the effectiveness of digital and online communications.

Online communications are perhaps the easiest to measure as web servers can indicate which and how many pages have been requested, the time spent on each page and even the type of computers that were used to request the page. However, this type of information is largely superficial and fails to provide insight into the user, their motivation to visit the site or the behavioural or attitudinal outcomes as a result of the interaction. Traditional measurement techniques of reach, frequency and target audience impressions are not readily transferrable to the Internet.

Review questions

1. If the process is difficult and the outcomes imprecise, why should organisations evaluate and monitor their marketing communications?
2. What are pre- and post-testing?
3. Write a brief report comparing recall and recognition tests.
4. What are the principal dimensions of likeability as a measure of advertising effectiveness?
5. Identify four ways in which sales promotions can be evaluated.

6. Write brief notes explaining why the use of media comparison techniques are insufficient when measuring the impact of public relations.

7. Why should the measurement of sales results be considered an inadequate measure of personal selling performance?

8. What are the techniques used to measure web site effectiveness? Are they any good?

9. Many organisations fail to undertake suitable research to measure the success of their campaigns. Why is this and what can be done to change this situation?

10. Comment on the view that, if a method of evaluation and testing lacks objectivity and testing, then the method should not be used.

MiniCase Measuring communication effectiveness at the Salvation Army: A bunch of well meaning amateurs or a leading-edge social relief organisation?

Jill Brown: University of Portsmouth

Founded in London, England, in 1865 by William and Catherine Booth, the Salvation Army is a distinctive and instantly recognisable Christian Church with its military uniform, bands and ranks. It is also one of the United Kingdom's largest charities.

As a global organisation, the Salvation Army operates throughout the world and delivers services ranging from community development, education, social and medical care and disease prevention among the world's poor, to international emergency services.

The Salvation Army seeks to communicate with a variety of diverse audiences. In the United Kingdom and the Republic of Ireland, there are the members (Salvationists) and clergy (officers) associated with around 700 churches, members of other Christian denominations, actual and potential donors (who may be individuals or businesses), grant-making foundations, journalists, local authorities and other social work funding bodies. In addition the Salvation Army also works at public affairs level, including working with MPs, because much of its work is related to the government's social welfare policies and with other relevant services including the medical profession and social services due to its pioneering work in the area of addictions and provision for the socially excluded.

In Britain, a major piece of marketing research undertaken in 1999 helped the Salvation Army to understand how it was perceived by its various audiences. Results showed that the organisation was respected and liked. It was seen as down to earth, action-oriented and honest but rather old-fashioned and amateur. The key images triggered by the brand 'Salvation Army' were of a homeless person being given a cup of soup, and people in military-type uniform playing in a band.

The challenge for this organisation is how to convey the diversity and professionalism of its worldwide social care activities and persuade people to look beyond its stereotypical image. Measuring the effectiveness of its communication activity is an essential part of this task.

Communications strategy

Following on from the perceptions research a communications strategy has been developed that aims to achieve a consistency of message when communicating with all the different audiences, and a vocabulary that emphasises a modern, professional, caring, results-driven organisation. In addition, a set of non-negotiable brand guidelines is being developed for churches to ensure greater consistency of logo, letter heading, name and signage. The Salvation Army has central control over fundraising, media relations and public affairs, but unlike the branches of a commercial organisation, the communications activities of its member churches are largely devolved to local level.

The Salvation Army's communication strategy is focused on three main areas: direct marketing, public and media relations and online communications.

Direct marketing

The Salvation Army's direct marketing activity reaches up to 25 million people a year and raises around £20 million. Its primary aim is fundraising but there are subsidiary objectives related to recruiting new donors

and raising awareness of the Salvation Army and its activities. Direct mail, door drops, magazine inserts, direct response television and radio advertising and online approaches are the main techniques used.

A specialist web media planner/buyer is employed to negotiate the best deals on a limited budget. Banner ads have been placed on Yahoo, television soap opera update websites and other sites frequented by the target donor audiences – typically people in their fifties and sixties. Pay per click deals were favoured until recently, but now a variety of deals are sought to give the best value for money in this rapidly evolving medium.

There is extensive measurement of direct marketing effectiveness including pretesting, marketing research and calculation of response rates and return on investment. Those who respond to the annual Christmas appeal for example, are sent a thank you letter that includes a questionnaire inquiring about their giving preferences and what further information they would like to receive from the Salvation Army. Focus groups composed of actual and potential supporters are used to check the likely effectiveness of direct mail messages or to test creative ideas for a poster and press advertising campaign.

Public and media relations

Good relations with the various media are very important to the Salvation Army but they are wary of making frequent, ill-judged public announcements. Instead, they seek to comment selectively on issues where they have expertise and something of value to say. Exposure in the local press has been found to be very effective and is linked with preparedness to give, so media awareness training is provided to church leaders and key people at local level to help them to deliver the Salvation Army's key messages to journalists.

Much of the Salvation Army's public relations activity is media-based but emphasis is also placed on personal influence. In addition, attendance at party political and other conferences, together with exhibitions is used to communicate with key opinion formers. The key objective here is 'to reinforce positive perceptions' according to Julius Wolff-Ingham, the Salvation Army's Head of Fundraising and Marketing.

The Salvation Army seeks to measure the effectiveness of its PR activity by participating in syndicated surveys undertaken by NFP Synergy, a marketing research agency that specialises in the not-for-profit sector. The quarterly awareness monitor asks a representative sample of the UK population questions such as, 'What are the top three UK charities which come to mind?', and 'Which charities work in the field of child protection?' The Salvation Army also under-takes surveys among two of its key audiences – MPs and journalists – to determine awareness levels and perceptions.

Online communications

The Salvation Army's web sites seek to serve a diverse range of people including donors, prayer supporters, school children studying the Salvation Army within the National Curriculum, potential service users, people seeking family members with whom they have lost contact, those exploring the Christian faith and looking for a Church, partner agencies, local authorities, journalists and professional practitioners.

There are separate web sites for young people (www.salvationarmy.org.uk/alove) and children (www.salvationarmy.org.uk/kids).

The current objectives for the web sites are:

(a) to emphasise that the Salvation Army is a world-wide evangelical Christian Church, as well as a human service provider, whose mission is to offer salvation in Christ to all;

(b) to offer relevant, accurate and up-to-date information about the Salvation Army in the United Kingdom to external enquirers, presented in an attractive, modern and user-friendly fashion;

(c) to provide news of Salvation Army activities and issues for an external audience;

(d) to reinforce the Salvation Army's identity and brand on the Internet in an attractive, dynamic, secure and consistent pattern, emphasising its contemporary relevance as well as its professional reliability;

(e) to provide an easy and inviting means for interested persons to contribute financially, in kind and as volunteers to the work of the Salvation Army in the United Kingdom, and so become and feel involved in its wider family;

(f) to create and serve an interactive online community of Salvationists and friends of the Salvation Army, encouraging them to welcome and support external visitors to the site.

The Salvation Army recognises the importance of developing a profile of its web site users. A recent survey of visitors to the youth web site sought to verify web analytics data that the primary users are youth leaders rather than young people themselves. Simple polls are regularly undertaken on the main web site to understand better who is visiting the site; Christian/non-Christian, views on ecology, moral issues, etc.

Web site performance is tracked primarily through analysing traffic data. The key metrics include number of visitors, page views, time on site, single-page bounces

and errors encountered. Referrals from search engines and other web sites are monitored to identify common keywords in order to optimise performance. Visitors are encouraged to use Web 2.0 tools such as Facebook and Stumbleupon to share the Salvation Army's web content virally, and the number of click-throughs from social networking web sites is increasing.

Cookies are used sparingly and do not persist for repeat visits (except for the sole purpose of preventing multiple voting in polls).

Constraints

Lack of money is a common reason for avoiding the evaluation of marketing communications, but in the not-for-profit sector where every penny is raised by voluntary donation, the issue is particularly acute. The Salvation Army operates under tight financial constraints but prides itself on reinvesting a very small amount of funds raised back into its fundraising activity.

It is constantly looking for cost-effective communication solutions. Even with major projects such as a web site relaunch, care is taken to spend donated money wisely and in-house focus groups are often used to verify web site useability. Staff numbers are low and so time resource is another constraint.

The well-meaning amateur image may persist but this is an organisation that understands the significance of measuring the effectiveness of its communication activity and does so rigorously.

MiniCase questions

1. Appraise the methods that the Salvation Army uses to monitor its corporate image.
2. To what extent do the Salvation Army's online communication evaluation techniques adequately measure the achievement of its web site objectives?
3. Suggest other ways in which the Salvation Army could measure the effectiveness of its promotional activity.
4. Should charitable organisations use donated funds to measure marketing communication effectiveness? Justify your response.

References

Armstrong, C. (1998) Sport sponsorship: a case study approach to measuring its effectiveness. *European Research*, **16**(2), 97–103.

Berthon, P., Pitt, L. and Watson, R. (1996) The world wide web as an advertising medium: toward an understanding of conversion efficiency. *Journal of Advertising Research*, **6**(1) (January/February), 43–53.

Biel, A.L. (1993) Ad research in the US. *Admap* (May), 27–9.

Briggs, R. and Hollis, N. (1997) Advertising on the web: is there response before click-through? *Journal of Advertising Research*, **37**(2), 33–46.

Burden, S. (2007) Case study: pre-testing mould-breaking ads. *Admap*, July/August, 48–9.

Campaign (2007) Cadbury 'gorilla' wins Campaign of the Year, *Campaign,* 13 December. Retrieved 16 January 2008 from http://www.brandrepublic.com/InDepth/Features/773064/Cadbury-gorilla-wins-Campaign-Year/.

Churchill, G.A., Ford, N.M. and Walker, C. (1990) *Sales Force Management.* Homewood, IL: Irwin.

Cox, S. (2007) Total recall: advertising exposure and engagement. *Admap*, 480, (February), 44–6.

Dichter, E. (1966) How word-of-mouth advertising works. *Harvard Business Review*, 44 (November/December), 147–66.

Doyle, P. and Saunders, J. (1985) The lead effect of marketing decisions. *Journal of Marketing Research*, **22**(1), 54–65.

Dreze, X. and Zurfryden, F. (1998) Is internet advertising ready for prime time? *Journal of Advertising Research* (May/June), 7–18.

Farrow, C. (1999) If it doesn't sell it isn't creative . . . true or false? *Marketing News* (October/November), 4–5.

Fendwick, P. (1996) The four ages of ad evaluation. *Admap* (April), 25–7.

Flesch, R. (1974*) The Art of Readable Writing*. New York: Harper & Row.

Goften, K. (1999) The measure of PR. *Campaign Report*, 2 April, 13.

Gordon, W. (1992) Ad pre-testing's hidden maps. *Admap* (June), 23–7.

Gray, R. (2000) The relentless rise of online research. *Marketing*, 18 May, 41.

Hauss, D. (1993) Measuring the impact of public relations. *Public Relations Journal* (February), 14–21.

Jones, J.P. (1990) Ad spending: maintaining market share. *Harvard Business Review* (January/February), 38–42.

Kitchen, P.J. (1993) Public relations: a rationale for its development and usage within UK fast-moving consumer goods firms. *European Journal of Marketing*, **27**(7), 53–75.

Krugman, H.E. (1988) Point of view: limits of attention to advertising. *Journal of Advertising Research*, **38**, 47–50.

Lannon, J. (2007) Marketing is the Boss, *Market Leader*, **39** (Winter). Retrieved 27 February 2008 from www.warc.com.

Lee, D.H. and Park, C.W. (2007) Conceptualization and measurement of multidimensionality of integrated marketing communications. *Journal of Advertising Research*, (September), 222–36.

Lodish, L.M. and Lubetkin, B. (1992) General truths? *Admap* (February), 9–15.

Marshall, D.W. and Cook, G. (1992) The corporate (sports) sponsor. *International Journal of Advertising*, **11**, 307–24.

Mazur, L. (1993) Qualified for success? *Marketing*, 23 January, 20–2.

Nielsen, A.C. (1993) Sales promotion and the information revolution. *Admap* (January), 80–5.

Oliver, G. (1990) *Marketing Today*, 3rd edn. Hemel Hempstead: Prentice-Hall.

Ovenden, M. (2007) 21st century media: right on target. *Marketing*, 5 December, 6–7.

Pearson, C.M. and Mitroff, I. (1993) From crisis prone to crisis prepared: a framework for crisis management. *Academy of Management Executive*, **7**(1), 48–59.

Peattie, K. and Peattie, S. (1993) Sales promotion: playing to win. *Journal of Marketing Management*, **9**, 255–69.

Quinn, J.B. and Mintzberg, H. (1992) *The Strategy Process*, 2nd edn. Englewood Cliffs, NJ: Prentice-Hall.

Reid, A. (2000) Testing Times, *Campaign,* 22 September, 40.

Robson, S. (2002) Group discussions. In *The International Handbook of Market Research Techniques* (ed. Robin Birn) Kogan Page, London.

Sherwood, P.K., Stevens, R.E. and Warren, W.E. (1989) Periodic or continuous tracking studies: matching methodology with objectives. *Market Intelligence and Planning*, **7**, 11–13.

Shultz, D.E. (1987) Above or below the line? Growth of sales promotion in the United States. *International Journal of Advertising*, **6**, 17–27.

Simms, J. (2007) Bridging the gap. *Marketing*, 12 December, 26–8.

Smit, E.G., van Meurs, L. and Neijens, P.C. (2006) Effects of advertising likeability: a 10-year perspective. *Journal of Advertising Research*, **46**(1) (March), 73–83.

Staverley, N.T. (1993) Is it right . . . will it work? *Admap* (May), 23–6.

Sturges, D.L., Carrell, B.J., Newsom, D.A. and Barrera, M. (1991) Crisis communication management: the public opinion node and its relationship to environmental nimbus. *SAM Advanced Management Journal*, (Summer), 22–7.

Thorncroft, A. (1996) Business arts sponsorship: arts face a harsh set of realities. *Financial Times*, 4 July, 1.

Vaughn, R. (1980) How advertising works: a planning model. *Journal of Advertising Research* (October), 27–33.

Watson, T. (1992) Evaluating PR effects. *Admap* (June), 28–30.

Wells, W., Burnett, J. and Moriarty, S. (1992) *Advertising: Principles and Practice*, 2nd edn. Englewood Cliffs, NJ: Prentice-Hall.

Wilson, K. (1993) Managing the industrial sales force of the 1990s. *Journal of Marketing Management*, **9**, 123–9.

Zielske, H.A. (1982) Does day-after recall penalise 'feeling' ads? *Journal of Advertising Research*, **22**(1), 19–22.