CHAPTER 28

New Developments in Best Practice Evaluation: Approaches, Frameworks, Models, and Methods

Jim Macnamara, University of Technology Sydney, Australia

Introduction

While evaluation in some fields of public communication in both the public and private sector has been in "stasis" (Gregory & Watson, 2008) or "deadlock" (Macnamara, 2015) with no consistent standards emerging despite 40 or more years of intensive focus (Likely & Watson, 2013), there have been a number of significant developments and advances recently that warrant close attention by scholars and practitioners. This chapter reviews the 'state of play' in evaluation of public sector communication and examines emerging developments and trends that inform the future with a particular focus on applications for public sector communication.

This is particularly relevant now, as public sector organizations come under increasing pressure for accountability and transparency as well as budget restraint – two of the key characteristics of public sector organizations identified by the editors in the introduction. For example, in 2016 the UK government announced 25–30 per cent cuts across most major government departments including those offering vital public services such as the Department of Health. The incoming government following the controversial decision by the UK to leave the European Community (EU) eased financial restraint somewhat, but maintained a tough stance on government spending and increased demands for accountability to citizens for expenditure. EU countries face tough budget pressures in the face of large-scale humanitarian immigration and economic recession faced by a number of member countries. Even in countries such as Australia that largely escaped the global financial crisis of 2008–2009, there is growing demand to rigorously evaluate government programs to justify as well as improve them.

In addition, the issue of values and expectations loom large in contemporary societies, with changing public attitudes towards politics and government. Whereas previous generations bestowed a great deal of faith in government, institutions, media, and traditional systems of politics such as political parties and voting, young citizens are increasingly sceptical and disengaged from traditional forms of political participation and citizenship (Bennett, 2008; Bennett, Wells, & Freelon, 2011; Carpentier, 2011; Coleman, 2013; Dalton, 2011). Traditional media, once mainstream channels for public sector communication, have declined substantially as sources of information for many citizens, particularly the young. In addition to increasingly relying on social media for news and information, many youth today are 'actualizing' rather than 'dutiful' citizens (Bennett et al., 2011; Schudson, 2003) engaged in new forms of "maximalist" and micro (grassroots) political expression and participation (Carpentier, 2011, pp. 17–18). These range from protest marches to major movements such as Occupy (Deluca, Lawson, & Sun, 2012) and the Umbrella Movement in Hong Kong (Jenkins, 2015). Societies also have become more diverse through multiculturalism and globalization. Public sector organizations are finding that they need to find new ways to communicate with and engage citizens, particularly those in marginalized and socially and culturally diverse communities. Therefore, they need to refine and expand their evaluation strategies, particularly in the context of the three-stage approach discussed here.

As noted by the editors in the introduction, public sector communication incorporates elements of and overlaps with a number of other practices including public relations (PR), public affairs, corporate communication, political communication, administration communication within public administration, as well as the evolving fields of government communication and what is broadly referred to as strategic communication and/or communication management. In examining new developments and future directions, this chapter will draw from all of those fields, arguing that there are significant benefits in taking a transdisciplinary approach.

The *Status Quo* of Evaluation

It is useful to briefly draw together other chapters in this section and summarize the status quo of public communication evaluation as a comparison point for discussing recent developments and future directions. Analysis of 40 years of discussion about evaluation of public communication in academic journal articles, books, and industry publications reveals

10 key factors and limitations shaping the evaluation landscape. It is important to call these out, as barriers and limitations need to be addressed before new ideas can take hold.

1. **Focus on outputs**

   Numerous studies have shown that evaluation in PR, corporate communication, strategic communication, and related fields of practice is predominantly focussed on outputs, rather than outcomes or impact of communication (Macnamara, Lwin, Adi, & Zerfass, 2015; Wright, Gaunt, Leggetter, Daniels, & Zerfass, 2009; Wright & Hinson, 2012; Zerfass, Verčič, Verhoeven, Moreno, & Tench, 2015). As discussed in the following, contemporary approaches shift the focus of evaluation to outcomes and impact.

2. **Assumptions about effects**

   Preoccupation with outputs with comparatively little critical attention paid to outcomes and impact is largely based on lingering assumptions about media effects. As is well-documented in media and communication literature, belief in strong and direct effects of mediated communication dominated thinking for much of the twentieth century influenced by transmissional models of communication (Berlo, 1960; Shannon & Weaver, 1949) and propaganda studies in the period of the two World Wars. As many scholars point out, direct effect theories of media and communication have been dismantled (e.g., Gauntlett, 2005) in favour of understandings of communication as transactional, contextual and contingent. Therefore, evaluation is essential, whereas it is not seen to be as important or even necessary at all when effects are assumed. Deeper knowledge about human communication and the challenges of attitude and behaviour change are key to avoiding assumptions and instead taking a social science approach.

3. **Focus on measurement vs. evaluation**

   Discussion in fields such as advertising, PR, government communication, strategic communication, and communication management is largely focussed on *measurement* rather than evaluation (e.g., Carroll & Stacks, 2004; CIPR, 2011; IAB, 2009; IPR, 2016; PRSA, 2014; Wright et al., 2009). Some articles, books, and manuals use both 'measurement' and 'evaluation' (e.g., Likely & Watson, 2013), but the key difference in these processes is often ignored. While measurement (the taking of measures) is necessary and a precursor to evidence-based evaluation, the latter involves "making a judgement" about the value or significance of something ("Evaluation", 2016). Evaluation is more specifically defined as "the process of gathering information about the merit or worth of a program for the purpose of making decisions about its effectiveness or for program improvement" (Owston, 2007, p. 606). Thus, measurement is only half of the process for assessing effectiveness and value. The models and frameworks discussed in this chapter highlight evaluation, not only measurement.

4. **Lack of formative, process, and summative approaches**

There is also a widespread misunderstanding of evaluation as limited to post-program assessment. While some researchers argue for four or even five stages of evaluation, most emphasize that there are at least two key stages of evaluation – *formative* and *summative* – or some prefer to apply three stages – *formative*, *process*, and *summative* evaluation (Sixsmith, Fox, Doyle, & Barry (2014; Valente, 2001; Valente & Kwan, 2013). Formative evaluation undertaken before programs are implemented is essential to understand existing audience awareness levels, perceptions, interests, needs, and preferred channels to inform planning, to provide benchmarks for later comparison, and for pre-testing messages and concepts. Process evaluation monitors progress and allows fine-tuning and adjustment of programs if necessary, while summative evaluation conducted after programs identifies outcomes and impact. Output oriented approaches focus on process evaluation, ignoring important steps such as pre-testing and failing to provide evidence of the effects of programs. The models and frameworks discussed in this chapter illustrate the importance of a progressive approach with considerable emphasis on formative evaluation.

5. **Excuses for not doing evaluation**

Lack of evaluation is justified in a number of ways. The three most commonly cited reasons for not doing or skimping on evaluation are (a) cost and lack of budget (Lindenmann, 2001; Wright et al., 2009); (b) lack of time (Wright et al., 2009); and (c) lack of demand by employers and clients (Baskin, Hahn, Seaman, & Reines, 2010). Lindenmann (2001) and others have pointed out that there are low-cost and even no-cost methods for evaluation and that evaluation is scalable from quick basic methods to rigorous social research (Macnamara, 1992, 2005, 2012). Recent mandating of evaluation by major public sector organizations such as the UK Cabinet Office (GCS, 2015) illustrates that there is demand. Other reasons cited for lack of evaluation are a lack of standards (Michaelson & Stacks, 2011; Ragan/NASDAQ OMX, 2013) and a search for a 'silver bullet' (Gregory & White, 2008; Likely & Watson, 2013). With standards in development and the search for a single solution widely dismissed by leading evaluation specialists such as Bauman and Nutbeam (2014) and Likely and Watson (2013), these so-called reasons for lack of evaluation are revealed to be largely excuses.

6. **Invalid and spurious methods**

The advertising and PR industries and some areas of corporate communication and communication management have been criticized for decades for the use of invalid and spurious methods of evaluation. The advertising industry has long relied on audience reach and recall of messages. Critics point out that simply reaching people with messages, and even recall of ads or messages, does not mean audiences were influenced by them. The PR industry has notoriously used so-called *advertising value equivalents* (AVEs), which are condemned as spurious and misleading on a number of grounds, as noted by Gregory in this section and many other authors (e.g., Lindenmann, 2003; Macnamara, 2000; Weiner & Bartholomew, 2006). In 2017 the International Association for Measurement and Evaluation of Communication (AMEC) launched a campaign to eradicate use of AVEs, which has been supported by professional organizations worldwide. The models and taxonomy of evaluation presented in this chapter identify a range of valid metrics and methods for evaluation.

7. **Media-centricity**

A further limitation and barrier to effective evaluation particularly prevalent in the advertising and PR industries is a preoccupation with media evaluation (Watson and Noble (2014). The International Association for Measurement and Evaluation of Communication (AMEC) was originally formed as the Association of Media Evaluation Companies, but soon realized that evaluation of communication involves much more than examining media content. But even the progressive evaluation of the UK Government Communication Service (GCS, 2015) continues to feature dashboards mainly showing media metrics based on content analysis and social media tracking. Contemporary models and frameworks for evaluation are designed to apply to a range of communication activities including Web communication, events, stakeholder and community engagement, and can even be applied to specialized public sector communication such as public consultation.

8. **Disciplinary siloes**

In tracing the history of evaluation of public communication and trying to establish standards, a US Task Force on Standardization of Communication Planning and Evaluation Models identified fragmentation in approaches to evaluation across practices such as advertising, PR, digital marketing, health communication, development communication, and other fields. Critical comparative analysis of literature showed that even closely related fields of practice are largely siloed and oblivious to knowledge, models, methods, and tools in other disciplines. Members of the task force called for a transdisciplinary approach and for the public communication sector to engage with fields such as program evaluation in public administration (Macnamara & Likely, 2017).

9. **Lack of research knowledge and skills**

In contrast with what are described here as 'excuses' for not doing evaluation, the preceding point and a number of other studies reveal that lack of knowledge and skills in relation to research is a key limitation in seeking better evaluation of public communication. In the first edition of their book on evaluation, Watson and Noble employed Dozier's (1992) categorization of communication practitioners as "technicians" focussed on production and outputs versus "managers" focussed on strategy and results to conclude that "it would appear that a generation of better-educated practitioners is needed to break the technician mould" (Watson & Noble, 2007, p. 46).

10. **Quantitative bias – the Modernist obsession with metrics**

A further limitation noted by this author in an address to the 2014 AMEC Summit on Measurement (the title is a further example of the focus on measurement rather than evaluation), is a dominance of quantitative thinking based on Modernist philosophies that privilege the 'scientific method' as well neoliberal tendencies to reduce all activities to metrics, especially financial metrics. Many researchers specializing in evaluation call for more qualitative research (e.g., Pawson & Tilley, 2001; Valente & Kwan, 2013). For example in discussing political communication, Karpf, Kreiss, and Neilsen recommend that to understand citizens' concerns, interests, needs, and views, in-depth qualitative methods of research are required such as "first-hand observation [ethnography], participation, and interviewing in the actual contexts where political communication occurs" (2014, p. 44). As shown in the following sections, contemporary approaches highlight a range of both quantitative and qualitative methods for evaluation.

Recent Developments and Advances

Against this background, a number of significant developments have taken place in recent years, some of which have been mentioned in preceding chapters, and some of which address the limitations described. These are briefly reviewed, particularly the most recent ones that offer hope for breakthroughs in the 'stasis' and 'deadlock' that has plagued evaluation research in public communication for decades.

The Barcelona Principles

A starting point for reform in evaluation of public communication for the public and private sectors was the declaration of the Barcelona Principles in 2010, so called because they were agreed by more than 200 delegates from 33 countries at an AMEC International Summit on Measurement in Barcelona. In many ways, the Barcelona Principles are basic 'home truths'. However, they were a tipping point for a series of further developments, including a major revision of the principles in 2015. Table 1 presents the Barcelona Principles as agreed in 2010 and the Barcelona Principles 2.0, which were jointly developed by AMEC; the International Communication Consultants Organization (ICCO); the Institute for Public Relations (IPR) in the US; the PR Consultants Association (PRCA) in the UK; the Public Relations Society of America (PRSA); and the Global Alliance for Public Relations and Communication Management[1] (AMEC, 2015).

Table 5.5.1. The original Barcelona Principles of 2010 and the Barcelona Principles 2.0 (AMEC, 2015).

	Barcelona Principles 2010	**Barcelona Principles 2.0 (2015)**
1.	Importance of goal setting and measurement	Goal setting and measurement are fundamental to communication and public relations
2.	Measuring the effect on outcomes is preferred to measuring outputs	Measuring communication outcomes is recommended versus only measuring outputs
3.	The effect on business results can and should be measured where possible	The effect on organizational performance can and should be measured where possible
4.	Media measurement requires quantity and quality	Measurement and evaluation require both qualitative and quantitative methods
5.	AVEs are not the value of public relations	AVEs are not the value of communication
6.	Social media can and should be measured	Social media can and should be measured consistently with other channels
7.	Transparency and replicability are paramount to sound measurement	Measurement and evaluation should be transparent, consistent and valid

The Barcelona Principles 2.0 represent a necessary improvement on the original principles in a number of respects including:

- Principle 1 is broadened beyond PR to 'communication';
- Principle 3 is broadened from "business results" to "organizational performance", which is more inclusive of the work of public sector and third sector organizations;
- Principle 4 is broadened beyond "media measurement"; and
- Replicability is removed from Principle 7 as this is applicable to quantitative research only.

However, there is still room for improvement in the Barcelona Principles 2.0. For instance:

- Principle 2 refers to "measuring communication outcomes", but does not specify evaluation of *impact* which is a stage beyond communication outcomes in most models. This will be further discussed under 'New evaluation frameworks and models';
- Principle 3 is broadened beyond a narrow focus on business, but it remains restricted to evaluating the effect on organizational performance and does not give any attention to impact on stakeholders or social impact;
- The terms 'measurement' and 'measuring' are used extensively and 'evaluation' is often not mentioned.

The 'march to standards'

Notwithstanding their weakness and generic nature, the Barcelona Principles provided a framework for what several industry commentators called the "march to standards" (Marklein & Paine, 2012). Two significant initiatives between 2010 and 2014 were as follows.

- In 2011 the Coalition for Public Relations Research Standards was established by AMEC, the IPR, and the Council of PR Firms (CPRF) to collaboratively develop standards for measurement and evaluation of PR within the framework of the Barcelona Principles. In 2012, the Coalition released *Proposed Interim Standards for Metrics in Traditional Media Analysis* (Eisenmann, Geddes, Paine, Pestana, Walton, & Weiner, 2012), which included definitions of key media content analysis terms such as 'items', 'impressions', 'mentions', 'tone', and 'sentiment', and described how these should be used;

- In 2012 the Conclave on Social Media Measurement Standards, known as the #SMMstandards Conclave or simply 'the Conclave' for short, was established. The Conclave involved collaboration by 11 professional communication organizations worldwide, as well as consultation with five media and advertising industry bodies and eight companies representing employer perspectives. . This was an important step in trying to achieve consistent terminology and compatibility of metrics across the public communication field.

At the fourth European Summit on Measurement in Dublin in 2012 the Coalition for Public Relations Research Standards o released three documents as the first stage of social media measurement standards: *Valid Metrics for Social Media* (Daniels, 2012) and *The Sources and Methods Transparency Table* and *Social Media Standard Definitions for Reach and Impressions* produced by the Conclave (2011) in consultation with the Digital Analytics Association. Subsequently, between 2011 and 2013 The Conclave developed standards for "content and sourcing", "reach and impressions", "engagement and conversation", "influence", "opinion and advocacy", and "influence and impact" (Conclave, 2011/2013).

Despite enthusiasm and much hard work by those involved, the 'standards' developed by these groups mainly comprise a set of definitions. These are a useful contribution, but do not comprise a set of standards for evaluation of public communication. Also, some so-called standards are superficial and contrary to established research literature. For example, engagement, which is a multidimensional concept described in organizational psychology as involving cognitive, affective, and conative dimensions (Macey & Schneider, 2008; Rhoades, Eisenberger, & Armeli, 2001), and as "two-way… give and take" between organizations and their stakeholders and publics (Taylor & Kent, 2014, p. 391), is described as including likes, comments, shares, retweets, and video views in social media. Furthermore, the initiatives reveal a continuing focus on media, on business outcomes, and *Americentrism*, with most members of these groups being Americans and all client organizations involved in these projects being US corporations.

The standards movement Mark II

A further initiative designed to move beyond definitions to more complete standards and to broaden focus internationally was the establishment in 2015 of the Task Force on Standardization of Communication Planning and Evaluation Models. Chaired by Canadian evaluation consultant Fraser Likely, the task force included academics and professional researchers who specialize in evaluation from Australia (including this author) and Europe as well as the US, along with members of the Institute for Public Relations (IPR) including some members of the IPR Measurement Commission (http://www.instituteforpr.org/ipr-measurement-commission). This task force attempted to synthesize myriad models of evaluation with a view to identifying approaches and methods that are theory-based and best practice, and therefore capable of being a standard.

Return to fundamentals

One of the first papers published by members of the Task Force on Standardization of Communication Planning and Evaluation Models already cited reviewed evaluation models published in PR and corporate communication literature from the early 1980s to the early 2000s compared with *program theory*, *program theory evaluation* (PTE), and *theory of change* models developed in the same period (Macnamara & Likely, 2017). The paper noted that some models of evaluation of PR and corporate, marketing, organizational, and government communication broadly followed the stages and processes of *program logic models*, but often modified or 'bastardized' these. For example, instead of the commonly used stages of *inputs*, *activities*, *outputs*, *outcomes*, and *impact* advocated in widely used models such as those of the Kellogg Foundation (1998/2004, 2010), United Way of America (Hatry, Houten, Plantz, & Greenway, 1996), and the University of Wisconsin Extension Program (UWEX) model (Taylor-Power & Henert, 2008), communication practitioners and evaluation service providers have created new terms such as 'outgrowths' and 'outflows'. A number of service providers also use 'black box' methods of evaluation based on proprietary algorithms. The paper argued for a focus on fundamental knowledge about evaluation as outlined in program theory (Weiss, 1972; Wholey, 1987), program theory evaluation (Rogers, Petrosino, Huebner, & Hacsi, T. (2000), theory of change (Anderson, 2005; Clark & Taplin, 2012), program logic models (Funnell & Rogers, 2011; Julian, 1997; Knowlton & Phillips, 2013; McLaughlin & Jordan, 1999), and other systematic, outcome and impact oriented approaches such as *realist evaluation* (Pawson & Tilley, 1997).

Also, Macnamara and Likely (2017) and other extended reviews of evaluation (e.g., Macnamara, 2018a) have criticized simplistic reductions of communication and information processing theory, such as illustrated in the AIDA (awareness, interest, desire, action) model still widely used in planning and evaluating advertising. Such models recognize and address only some of the steps in communication identified by communication theorists and social psychologists such as W. J. McGuire, who initially identified six steps (McGuire, 1985), but later expanded this to 13 steps (McGuire, 2001).

New Evaluation Frameworks and Models

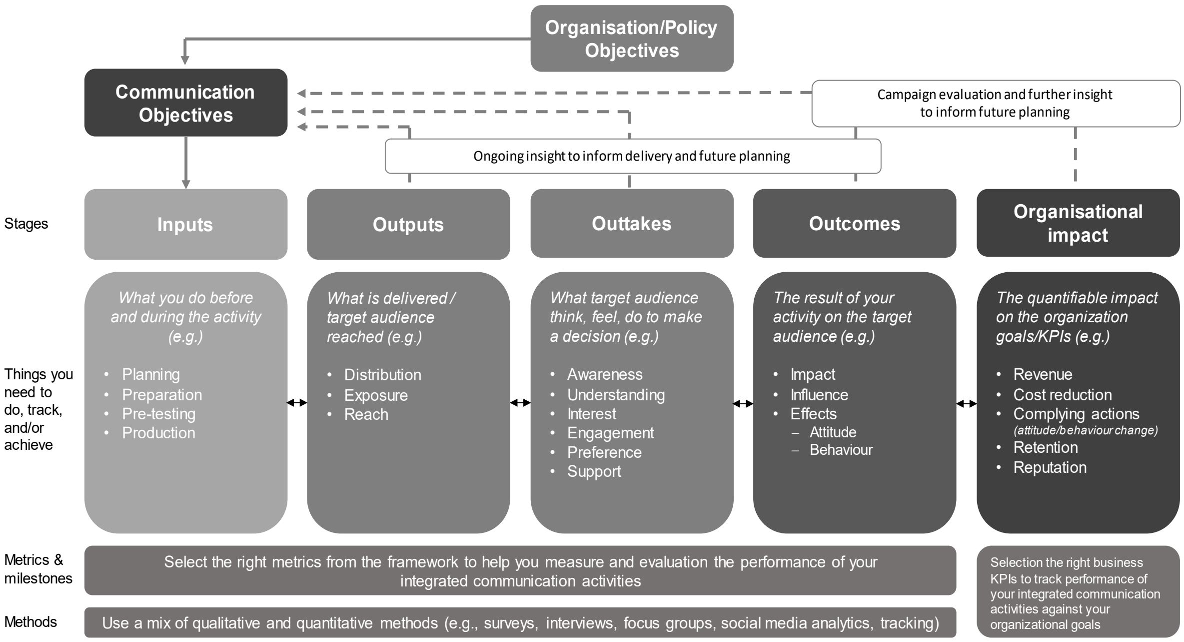
Since 2015 a number of new frameworks and models for evaluation of public communication in and by public as well as the private sector organizations have been developed based on transdisciplinary knowledge drawn from fields such as program evaluation and other systematic approaches (Rossi, Lipsey, & Freeman, 2004; Wholey, Hatry, & Newcomer, 2010). These offer insights into best practice as well as application of theory-based evaluation and, not insignificantly in the context of this text, a number of these initiatives are in the public sector.

The UK Cabinet Office evaluation framework

Following considerable focus on evaluation of communication within the EU government (European Commission, 2014; Henningsen, Traverse Healy, Gregory, Johannsen, Allison, Bozeat, et al., 2014), the UK Government Communication Service (GCS) mandated evaluation for all UK government departments and agencies referred to in Britain as arm's length bodies (ALBs) in 2014 and set quite demanding standards for reporting. Two supporting initiatives gave credibility and validity to the GCS initiatives. First, it established an Evaluation Council made up of independent consultants, professional researchers, and academics with the role of advising GCS and the Cabinet Office on evaluation of public communication and reviewing proposals for major campaigns. Second, GCS introduced a major professional development program for government communicators across the UK civil service. This offers workshops and online materials to advance staff from an introduction to systematic evaluation to 'evaluation champions'. The professional development program is supported and guided by *GCS Evaluation Capability Standards* based on surveys of communication staff to monitor and improve knowledge and skills.

Through the advice of its evaluation council and also direct consulting with academic researchers, the UK GCS evaluation framework introduced in 2015 and updated in 2016 (GCS, 2016) and again in 2018 in the *GCS Evaluation Framework 2.0* (GCS, 2018), reflects the key concepts and principles of program logic models and communication and information processing theory such as that of McGuire (2001). The over-arching model of evaluation, which is part of the framework (GCS, 2016, 2018), retains the stage of *outtakes* introduced to PR evaluation models by Fairchild (1997) in the UK and then Lindenmann (2003) in the US (see Figure 5.5.1). However, importantly, it shows the process of planning and evaluating programs beginning with communication objectives and these being clearly linked to organizational objectives. The processes of evaluation of UK government communication are based on a five-stage program logic model, with findings from formative, process, and summative evaluation used as feedback to fine-tune and adjust programs if necessary (shown in the dotted lines and arrows).

Figure 5.5.1. The UK Government Communication Service evaluation framework (GCS, 2016, 2018).



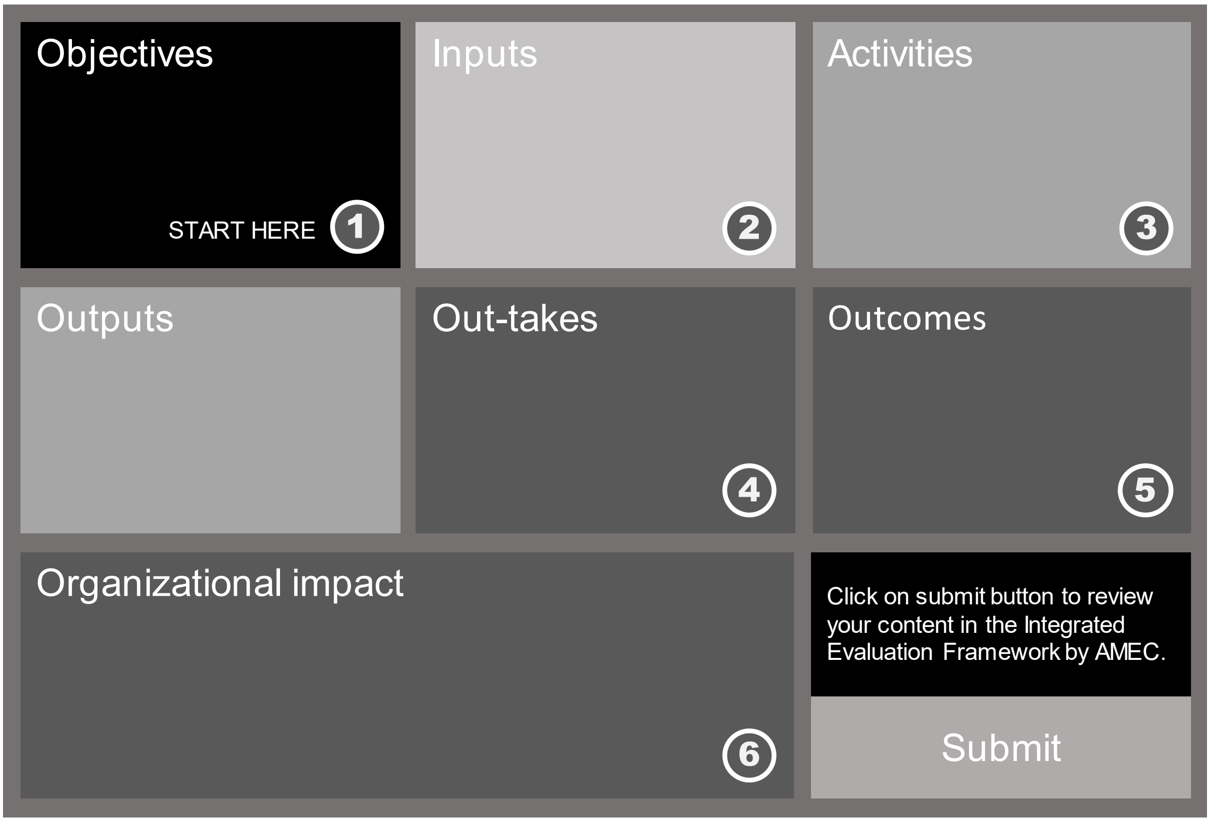
Many of the traditional barriers to evaluation referred to previously in this chapter, such as using invalid methods, lack of formative evaluation, and excuses were overcome in the UK GCS through the roll-out of extensive training in evaluation as part of the GCS professional development program, the cultivation of 'evaluation champions', and also a senior management decision to mandate evaluation. Major campaigns need to be submitted to the GCS Evaluation Council and are not approved unless rigorous evaluation is included.

The AMEC integrated evaluation framework

In mid-2016, the International Association for Measurement and Evaluation of Communication (AMEC) launched a major revision of its former *Valid Metrics Framework*. This also uses a program logic model approach, albeit it uses a six-stage model – inputs, activities, outputs, outtakes, outcomes, and impact. In other words, it uses a classic five-stage program logic model as used in other fields, but adds 'outtakes' as PR-oriented evaluators frequently do. Whether practitioners will be able to differentiate outtakes from outcomes is a question that only time will tell. It may only add to the confusion discussed in the following section. Nevertheless, the AMEC Integrated Evaluation Framework (AMEC, 2016) was developed with considerable input from AMEC's Academic Advisory Group, of which this author served as chair, as well as a committee of evaluation specialists led by Richard Bagnall, then CEO of Prime Research in the UK. It takes its name from its integration of evaluation theory and academic research, existing industry matrices and models, and its capability to evaluate integrated communication involving paid, earned, shared, and owned channels.

Some features do particularly set the AMEC Integrated Evaluation Framework apart and advance the field of practice considerably. The first is that the AMEC framework is an interactive online tool, not a static diagram (http://amecorg.com/amecframework). This means that users can input data into the various steps and stages, including organization objectives and communication objectives, and then progressively add inputs, outputs, and so on. Thus, it is a working tool. A second key feature of the AMEC framework that advances evaluation considerably for communication practitioners is that the framework is supported by a range of resources to guide practitioners through the process of evaluation. These include guidelines for setting SMART objectives; links to the *Dictionary of Public Relations Measurement and Research* (Stacks & Bowen, 2013); and a number of downloadable case studies that provide samples of the framework in use. One of the key resources in terms of using the framework is a taxonomy of evaluation as discussed in the following section.

Figure 5.5.2. The interface of the AMEC Integrated Evaluation Framework interactive online tool (AMEC, 2017).

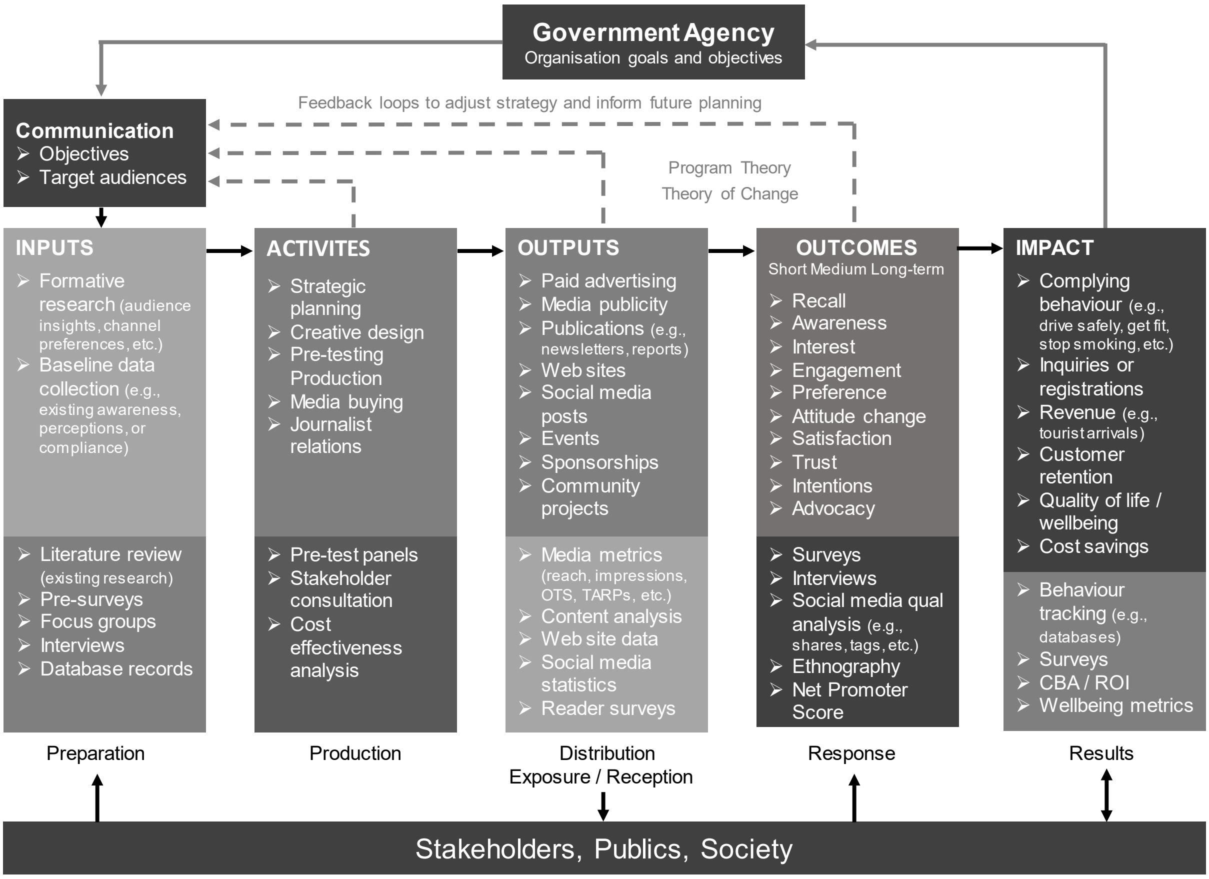


The integrated model of evaluation evolves

In 2016, the Strategic Communications Branch of the New South Wales state government in Australia adopted an evaluation model as part of a research project to update its framework for evaluation of advertising and other forms of public communication (DPC, 2016). This reverted to a classic five-stage program logic model, but drew on the AMEC integrated evaluation framework and the GCS model in terms of identifying appropriate metrics and methods for evaluation and applying these to public sector communication.

In addition, the NSW Government evaluation model for advertising and communication broke new ground by being the first evaluation model to explicitly represent stakeholders, publics and society in the communication process and to identify two-way flow of information, feedback and impact represented by arrows in the model (see Figure 5.5.3). Previous models represented theory of change and influence as flowing top-down from organizations to others who were conceptualized only as 'target' audiences. Thus, they reflected a one-way, organization-centric approach to evaluation. However, this model continued to identify communication objectives as being unilaterally determined by the organization.

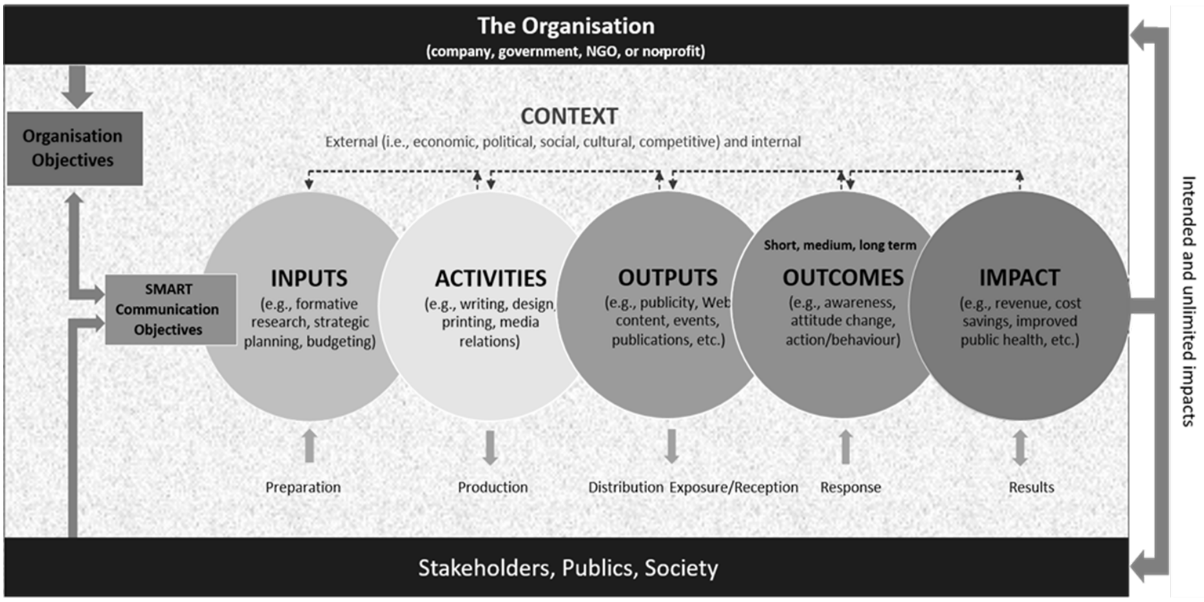
Figure 5.5.3. Evaluation model for advertising and communication adopted by the Strategic Communications Branch of the New South Wales Department of Premier and Cabinet (DPC, 2016).



The research that informed the NSW Government evaluation model for communication continued for a book reviewing evaluation of advertising, PR and specialist fields such as health communication internationally (Macnamara, 2018a), which produced a further iteration of an integrated evaluation model. The Public Relations Institute of Australia adopted an early version of this evolving integrated model of evaluation (PRIA, 2017), that was later published in Macnamara (2018a) and Macnamara 2018b, p. 192), so named because it attempts to combine the best features of other models and address key deficiencies as explained in the following.

As well as being based on the five stages of classic program logic models and representing these as over-lapping stages rather than separate 'boxes' (see Figure 5.5.4), there are three important features of the integrated evaluation model that are particularly applicable to public sector organizations.

Figure 5.5.4. The integrated evaluation model (Macnamara, 2018b, p. 192), an early version of which was adopted by the PRIA (2017).



First, in addition to linking outcomes and impact to government or agency objectives, it recognizes that communication objectives and inputs to planning should take into account the views, needs, and interests of stakeholders, publics and society generally – not only those of the organization, as reflected in other models. In short, evaluation begins even before communication objectives are set and formative evaluation continues during the input stage to inform planning.

Second, it even more explicitly shows that, while the activities and outputs stages of communication involve information flow from the organization to stakeholders, publics and society, outcomes should be evaluated based on response and reactions from stakeholders and publics (represented by arrows below each stage in the model).

Third, as first proposed in the NSW Government communication evaluation model, impact should be evaluated from the perspective of both the organization and stakeholders, publics and society. Recognition of the need to evaluate outcomes and impact from the perspective of stakeholders and publics as well as the organization aligns with program evaluation theory and program logic models (Kellogg Foundation, 1998/2004; Taylor-Power & Henert, 2008; Wholey, Hatry, & Newcomer, 2010) and with Excellence theory of PR, which calls for evaluation to be conducted at (a) program level; (2) functional level (e.g., department or unit); (3) organizational level; and (4) societal level (L. Grunig, J. Grunig & Dozier, 2002, pp. 91–92). This is an important consideration in evaluation by public sector organizations, particularly in democratic societies in which government must be responsive and responsible to citizens and society.

Two other important characteristics of this latest integrated evaluation model are that it specifies that *unintended* as well as *intended* impacts should be considered in evaluation – a factor overlooked in other models – and uniquely it identifies that *context* is an important factor in evaluation. The external economic, political, social, cultural and competitive context, as well as the internal context within an organization, significantly affect what can be achieved in communication and must be taken into consideration. Thus, this integrated evaluation model presents a more comprehensive and holistic overview of the key concepts, principles and processes of evaluating public sector communication.

A Taxonomy of Evaluation

While models provide useful overviews of processes, there is still the question of what is done when in terms of doing evaluation. In their leading PR textbook, Cutlip, Center and Broom note repeatedly in editions from 1985 to the late-2000s that "the common error in program evaluation is substituting measures from one level for those at another level" (1985, p. 295; 1994, p. 414; Broom, 2009, p. 358). This warning has been echoed by emeritus professor of public relations Jim Grunig who said that many practitioners use "a metric gathered at one level of analysis to show an outcome at a higher level of analysis" (2008, p. 89). More broadly, the UWEX guide to program evaluation similarly says that "people often struggle with the difference between outputs and outcomes" (Taylor-Power & Henert, 2008, p. 19).

In working with the Task Force on Standardization of Communication Planning and Evaluation Models and the AMEC team developing its integrated evaluation framework, a key step was synthesizing a wide range of literature that identifies the inputs, activities, outputs, outcomes, and impact of public communication and the methods applicable to evaluating them. This enabled production of a taxonomy of evaluation for communication.

The term taxonomy is used in preference to typology as, even though the terms are often used interchangeably, a taxonomy categorizes empirical entities based on evidence, whereas a typology is typically a conceptual construct (Bailey, 1994, p. 6). The taxonomy of evaluation of communication identifies four levels in each of up to six stages of communication (inputs, activities, outputs, outtakes, outcomes, and impact. Under each stage (the *macro* level); the taxonomy lists key steps required in that stage (*meso* level); then lists typical milestones and metrics for showing achievement of those steps (*micro* level); and then lists typical methods to demonstrate milestones or generate metrics required for that stage.

By adding this detail, frameworks evolve towards a working model of evaluation for public communication. This taxonomy also clearly draws attention to the fact that outputs are barely half way to achieving outcomes and impact. It points to the need to go beyond distribution, exposure, and even reception of information by audiences to gaining *attention*; creating *awareness* and *understanding*; generating *interest* or *liking*; creating *engagement* and *participation* and, even further, *consideration*, which AMEC and some other models refer to as outtakes and some refers to as short-term outcomes. Furthermore, based on McGuire's (2001) steps of information processing, communication needs to go even further, sometimes to *learning* or creating new *knowledge* in audiences, *attitude change*, *satisfaction*, *trust*, *preference*, *intention* (e.g., to buy or act), and *advocacy* (urging others to buy or act). Ultimately, impact is identified in terms of *complying action* or other results such as positive *reputation*, *relationships*, *organization change*, or *public/social change*. Key milestones and metrics as well as typical methods to generate these metrics and demonstrate milestones are also listed for each of these stages and steps.

The taxonomy shown in Table 5.5.2 arranges key steps, milestones and metrics, and evaluation methods in six stages to match the AMEC framework (Macnamara, 2016). However, a complete version of this taxonomy similarly arranges key steps, milestones and metrics, and methods for five-stage, four-stage, and even three-stage (inputs, outputs, outcomes) logic models. In a five-stage framework, outtakes are reclassified as 'short-term outcomes' in line with classic logic models, with outcomes being 'intermediate' and 'long-term outcomes'. In a four-stage framework (inputs, outputs, outcomes, and impact), outtakes are arranged as above (i.e., as short-term outcomes) and inputs are combined with activities. In a three-stage framework (inputs, outputs, outcomes), impact is renamed 'long-term outcomes', as well as the other combined stages listed above (Macnamara, 2018a, 2018b).

Table 5.5.2. A taxonomy of communication evaluation stages, key steps, milestones and metrics, and methods (Macnamara, 2016).

STAGES *Macro-level*	**INPUTS**	**ACTIVITIES**	**OUTPUTS**	**OUTTAKES** Short-term outcomes	**OUTCOMES** Intermediate	**IMPACT** Long-term
Short definition	*What you need in preparation for communication*	*Things you do to plan and produce your communication*	*What you put out that is received by target audiences*	*What audiences do with and take out of your communication*	*Effects that your communication has on audiences*	*The results that are caused, in full or in part, by your communication*[2]
KEY STEPS *Meso-level*	• **Objectives** • **Budget** • **Resources** *(e.g., staff, agencies, facilities, partnerships)*	• **Formative research** • **Planning**[3] • **Production** *(e.g., design, writing, media buying, media relations, media partnerships, etc.)*	• **Distribution** • **Exposure** • **Reception**[4]	• **Attention** • **Awareness** • **Understanding** • **Interest / liking** • **Engagement** • **Participation** • **Consideration**	• **Learning / knowledge**[5] • **Attitude change** • **Satisfaction** • **Trust** • **Preference** • **Intention** • **Advocacy**	• **Compliance / complying actions** • **Reputation** • **Relationships** • **Organisation change** • **Public/social change**
Example METRICS & MILESTONES *Micro-level*	• SMART objectives • Benchmarks / baseline data • Targets / KPIs	• Baselines / benchmarks *(e.g., current awareness)* • Audience needs, preferences, etc. • Strategic plan • Evaluation plan • Pre-test data *(e.g., creative concepts)* • Content produced *(e.g., media releases, Websites)* • Media relations	• Advertising TARPs • Audience reach • Impressions/OTS[6] • CPM • Clickthroughs • Publicity volume • Share of voice • Tone/sentiment/ favourability • Messages placed • Posts, tweets, etc. • E-marketing volume • Event attendance	• Unique visitors • Views • Response *(e.g., follows, likes, tags, shares, retweets)* • Return visits/views • Recall (unaided, aided) • Positive comments • Positive response in surveys, etc. • Subscribers *(e.g., RSS, newsletters)* • Inquiries	• Message acceptance • Trust levels • Statements of support or intent • Leads • Registrations *(e.g., organ donor list)* • Brand preference • Trialling • Joining • Reaffirming *(e.g., staff satisfaction)*	• Public/s support • Meet targets *(e.g., blood donations; cancer screening membership, etc.)* • Sales increase • Donations increase • Cost savings • Staff retention • Customer retention/loyalty • Quality of life / wellbeing increase
METHODS of evaluation	• Internal analysis • Environmental scanning • Feasibility analysis • Risk analysis	• Metadata analysis *(e.g., past research and metrics)* • Market/audience research *(e.g., surveys, focus groups, interviews)* • Stakeholder consultation • Case studies *(e.g., best practice)* • SWOT analysis (or PEST, PESTLE, etc.) • Pre-testing panels • Peer review / expert review	• Media metrics *(e.g., audience statistics, impressions, CPM)* • Media monitoring • Media content analysis *(quant)* • Media content analysis *(qual)* • Social media analysis *(quant and qual)* • Activity reports *(e.g., events, sponsorships)*	• Web statistics *(e.g., views, downloads)* • Social media analysis *(qual – e.g.., comments)* • Feedback *(e.g., comments, letters)* • Ethnography *(observation)* • Netnography *(online ethnography)* • Audience surveys *(e.g., re awareness, understanding, interest)* • Focus group *(as above)* • Interviews *(as above)*	• Social media analysis *(qual)* • Database statistics *(e.g., identifying sources)* • Ethnography *(observation)* • Netnography *(online ethnography)* • Opinion polls • Stakeholder surveys *(e.g., re satisfaction, trust)* • Focus groups *(as above)* • Interviews *(as above)* • Net Promoter Score (NPS)[7]	• Database records *(e.g., blood donations, health outcomes, membership)* • Sales tracking • Donation tracking • CRM data • Staff survey data • Reputation studies • Cost Benefit Analysis • ROI *(if there are financial objectives)* • Econometrics[8] • Quality of life scales & wellbeing measures

These options recognize that there is unlikely to be a single evaluation framework or model. Nor is it possible to have a single evaluation method, given that different communication activities and campaigns have different objectives. Rather, most industries and sectors establish *standards* (plural) – a range of practices that conform to key principles that are agreed and formalized in manuals, guides, and other publications.

In 2018 the Directorate-General for Communication (DG COMM) of the European Commission began to update its tools for evaluation of communication including a list of indicators for evaluating events, publications, media relations and other communication activities (European Commission, 2018), which were adapted from the AMEC taxonomy (Macnamara, 2016) and more recent versions (Macnamara, 2018a, pp. 121–131). However, DG COMM communicators are bound by the *Better Regulation Guidelines* of the European Commission (2017), which specify the processes for evaluation of all activities, not only communication, and contain a model that focusses on evaluation of "outputs, results and impacts" (European Commission, 2017, p. 58). This illustrates the continuing use of different terminology and models for evaluation of public sector communication.

The work of AMEC, the IPR Measurement Commission, the UK Government Communication Service, the European Commission Directorate-General for Communication, the NSW Department of Premier and Cabinet Strategic Communications Branch, and academic researchers are likely to be ongoing for some time. However, the model, taxonomies, indicators and guidelines summarized here represent significant advances in thinking about and applying evaluation.

The *Quo Vadis* of Evaluation

Having started by summarizing the *status quo* over the past few decades, and then looking at recent developments, it is appropriate to end by asking the question *quo vadis* (Latin for 'where are you going?).

Some of the important future developments will be fine-tuning and evolution of frameworks and models such as that of AMEC and those used by government in the Australia, the EU and the UK. Greater use of qualitative research is also recommended by a number of evaluation specialists (Pawson & Tilley, 2001; Valente & Kwan, 2013). In addition, there are a number of sophisticated new methods designed to cope with changing societal conditions, expectations, and cultural practices. Two examples are briefly introduced, further illustrating the opportunities available through transdisciplinary approaches to draw on the methods and expertise of other fields.

Behavioural insights

Behavioural insights, also referred to as behavioural economics, is a research method that draws on social psychology and economics to "explain why people behave in ways that deviate from rationality as defined by classical economics" (Marteau, Ogilvie, Roland, Suhrcke, & Kelly, 201, p. 223). The focus of behavioural insights is understanding the influences – triggers, if you will – that shape people's choices in relation to certain behaviours and then manipulating those influences to create the desired behaviours (Thaler & Sunstein, 2008, p. 6). Because behavioural insights are gained in order to stimulate desired behaviours in people, the field has become known colloquially as *nudge* communication, also referred to as *nudge* marketing, a term created by the pioneers of this field of practice, Richard Thaler and Cass Sunstein (2008) in their book *Nudge: Improving Decisions about Health, Wealth, and Happiness*.

Several governments have adopted behavioural economics / behavioural insights including the UK Government, which pioneered application of Thaler and Sunstein's concept in setting up a behavioural insights team in the Cabinet Office in 2010 before spinning it off as a social purpose company (http://www.behaviouralinsights.co.uk) headed by British psychologist David Halpern. The Institute for Government in the UK also has advocated the application of behavioural insights, although its focus has been mainly on the development and implementation of public policy rather than evaluation of communication, as highlighted in a 2010 report *MINDSPACE: Influencing Behaviour Through Public Policy* (Dolan, Hallsworth, Halpern, King, & Vlaev, 2010). In the US, Harvard University's John F. Kennedy School of Government has established the Behavioural Insights Group (BIG), and the White House set up a Nudge Unit in 2014 (Nesterak, 2014). In Australia the state government of New South Wales has established a Behavioural Insights Community of Practice (http://bi.dpc.nsw.gov.au) to share knowledge across departments and agencies.

Behavioural insights inform evaluation by identifying the 'triggers' to attitude and behaviour change. However, given that these include strategic use of incentives and appeals to emotions, ego, subconscious cues, norms, and default behaviours, as well as restrictions such as regulation and legislation (see Table 5.5.3), it is important that behavioural insights are applied ethically. Thaler and Sunstein believe that it is appropriate for governments and public sector organizations to use behavioural insights and nudge techniques in a spirit of what they call *libertarian paternalism* (Sunstein & Thaler, 2003, p. 1160).

Table 5.5.3. The MINDSPACE checklist of influences on human behaviour.

Influence	Description
Messenger	We are heavily influenced by who communicates information
Incentives	Our responses to incentives are shaped by predictable mental shortcuts such as strongly avoiding losses
Norms	We are strongly influenced by what others do
Defaults	We ‘go with the flow' of pre-set options
Salience	Our attention is drawn to what is novel and seems relevant to us
Priming	Our acts are often influenced by subconscious cues
Affect	Our emotional associations can powerfully shape our actions
Commitments	We seek to be consistent with our public promises, and reciprocate acts
Ego	We act in ways that make us feel better about ourselves

Source: Dolan, Hallsworth, Halpern, King, & Vlaev (2010).

Behavioural insights-based approaches are also relevant to evaluation because they involve continual evaluation. Every step in a program aimed at changing behaviour is evaluated and compared with other alternative approaches to achieve the objective and only the most effective activities are continued. See Thaler and Sunstein (2008) for more information.

Sense making methodology

Another recently developed and advanced method of planning, implementing, and evaluating public communication is *sense making methodology* (SSM). There are four key features of sense making methodology that are pertinent for research including evaluation as follows.

1. SMM avoids categorizing people in advance of research, such as by demographics or psychographics, instead allowing them to speak for themselves and describe who they are and how they feel (referred to in SSM as *verbing* rather than *nouning* because it focuses on what people do and say rather than naming them and putting labels on them).

In the words of the developers of SMM, participants in SMM research are "theorists of their own worlds" (Dervin & Foreman-Wernet, 2013, p. 158).

2. SMM uses more qualitative than quantitative research to probe deeply into people's perceptions, attitudes, interests, desires, fears, beliefs, and so on.

3. In implementing research SMM allows participants time for reflection. SMM typically uses focus groups and discussion forums. In these a particular SMM technique is *sense making journaling* in which participants are asked to write down their reactions to subjects being discussed, reflect on them, and prepare their thoughts before answering questions or expressing their views. Dervin and Foreman-Wernet say that traditional interviewing and surveys put people on the spot to give an answer, whereas in journaling "space is opened for what is usually left unsaid" (2013, p. 155).

4. A fourth key principle that makes SSM fundamentally different to other top-down and expert led approaches to research and evaluation is that the methodology is based on an understanding and acceptance that "both organizations and constituencies have expertise to share, common struggles to ponder, and capacities to teach and learn from each other" (Dervin & Foreman-Wernet, 2013, p. 160). Dervin and her co-author state frankly: "in SMM public communication is defined as the means to not merely change constituencies but to change organizations" (p. 160).

Conclusions

Effective evaluation to inform future planning and strategy as well as provide accountable, transparent reporting is available, but requires: (1) increased focus on outcomes and impact of communication, including on stakeholders, publics and society as well as the organization; (2) recognition and implementation of formative, process, and summative evaluation; (3) greater use of qualitative research; (4) adoption of contemporary frameworks and models based on program evaluation theory and specifying systematic methods of evaluation; and (5) considering innovative approaches such as SMM and behavioural insights.

The bi-directional flow of feedback and influence reflected in the integrated evaluation model shown in Figure 5.5.4 supports the above shifts and approaches.

In addition, a key to implementation is increasing knowledge and skills among practitioners to deploy these strategies. With responsibilities to communicate with citizens effectively and use public funds wisely, public sector communicators should be at the forefront of these developments.

References

AMEC (Association for Measurement and Evaluation of Communication). (2015). Barcelona principles 2.0. London, UK: Author. Retrieved from http://amecorg.com/barcelona-principles-2-0

AMEC (Association for Measurement and Evaluation of Communication). (2017). Integrated evaluation framework. London, UK: Author. Retrieved from http://amecorg.com/amecframework

Anderson, A. (2005). *The community builder's approach to theory of change: A practical guide to theory and development*. New York, NY: The Aspen Institute Roundtable on Community Change.

Bailey, K. (1994). *Typologies and taxonomies: An introduction to classification techniques*. Thousand Oaks, CA: Sage.

Baskin, O., Hahn, J., Seaman, S., & Reines, D. (2010). Perceived effectiveness and implementation of public relations measurement and evaluation tools among European providers and consumers of PR services. *Public Relations Review, 36*, 105–111.

Bauman, A., & Nutbeam, D. (2014). *Evaluation in a nutshell: A practical guide to the evaluation of health promotion programs* (2nd ed.). North Ryde, NSW: McGraw-Hill.

Behavioural Insights. (2015). Web site resources. London, UK. Retrieved from http://www.behaviouralinsights.co.uk

Bennett, W. (2008). Changing citizenship in the digital age. In W. Bennett (ed.), *Civic life online* (pp. 1-24). Cambridge, MA: MIT Press.

Bennett, W., Wells, C., & Freelon, D. (2011). Communicating civic engagement: Contrasting models of citizenship in the youth Web culture. *Journal of Communication, 61*(5), October, 835–856. doi:10.1111/j.1460-2466.2011.01588.x

Berlo, D. (1960). *The process of communication: An introduction to theory and practice*. New York, NY: Harcourt/Holt, Rinehart & Winston.

Carpentier, N. (2011). *Media and participation: A site of ideological-democratic struggle*. Chicago, IL: Intellect/University of Chicago Press.

Carroll, T., & Stacks, D. (2004). Bibliography of public relations measurement. Gainesville, FL: Institute for Public Relations. Retrieved from http://www.instituteforpr.org/wp-content/uploads/PRM_Bibliography.pdf

Chen, H. (1990) *Theory-driven Evaluation.* Thousand Oaks, CA: Sage.

Chen, H., & Rossi, P. (1983). Evaluating with sense: The theory-driven approach. *Evaluation Review, 7*, 283–302.

CIPR (Chartered Institute for Public Relations). (2011). Research, planning and measurement toolkit. Retrieved from http://www.cipr.co.uk/sites/default/files/Non-member%20excerpt%202_0.pdf

Clark, H., & Taplin, D. (2012). *Theory of change basics: A primer on theory of change*. New York, NY: Actknowledge.

Coleman, S. (2013). *How voters feel.* New York, NY: Cambridge University Press.

Conclave on Social Media Measurement Standards. (2011/2013). The standards [Web site]. Retrieved from http://smmstandards.wixsite.com/smmstandards

Dalton, R. (2011). *Engaging youth in politics: Debating democracy's future*. New York: Idebate Press.

Daniels, M. (2012, June). Valid metrics framework. Presentation to the 4th European Summit on Measurement, Dublin, Ireland. Retrieved from http://amecorg.com/downloads/dublin2012/Valid-Metrics-Framework-Mike-Daniels.pdf

Deluca, K., Lawson, S., & Sun, Y. (2012). Occupy Wall Street on the public screens of social media: The many framings of the birth of a protest movement. *Communication, Culture & Critique, 5*(4), 483–509.

Dervin, B., & Foreman-Wernet, L. (2013). Sense-making methodology as an approach to understanding and designing for campaign audiences. In R. Rice & C. Atkin (Eds.), *Public communication campaigns* (4th ed., pp. 147–162). Thousand Oaks, CA: Sage.

Dolan, P., Hallsworth, M., Halpern, D., King, D., & Vlaev, I. (2010). *MINDSPACE: Influencing behaviour through public policy.* London, UK: Institute for Government. Retrieved from www.instituteforgovernment.org.uk/our-work/better-policy-making/mindspace-behavioural-economics

Dozier, D. (1992). The organizational roles of communications and public relations practitioners. In J. Grunig (Ed.), *Excellence in public relations and communication management* (pp. 327–356). Hillsdale, NJ: Lawrence Erlbaum.

DPC (Department of Premier and Cabinet). (2016). NSW government evaluation framework for advertising and communication. Sydney, NSW: NSW Government.

Eisenmann, M., Geddes, D., Paine, K., Pestana, R., Walton, F. & Weiner, M. (2012). Proposed interim standards for metrics in traditional media analysis. Gainesville, FL: Institute for Public Relations. Retrieved from http://www.instituteforpr.org/topics/proposed-interim-standards-for-metrics-in-traditional-media-analysis

European Commission. (2014, August). *Measuring the European Commission's communication: Technical and methodological report*. Brussels, Belgium: Author. Retrieved from http://ec.europa.eu/dgs/communication/about/evaluation/index_en.htm

European Commission. (2017). Better regulation guidelines. Brussels, Belgium: Author. Retrieved from https://ec.europa.eu/info/sites/info/files/better-regulation-guidelines.pdf

European Commission. (2018). Communication network indicators. Brussels, Belgium: Author.
Evaluation. (2016). Merriam-Webster dictionary. Retrieved from http://www.merriam-webster.com/dictionary/evaluate
Fairchild, M. (1997). *How to get Real value from public relations*. London, UK: ICO.
Funnell, S., Rogers, P. (2011). *Purposeful program theory: Effective use of theory of change and logic models*. San Francisco, CA: Jossey-Bass.
Gauntlett, D. (2005). *Moving experiences: Media effects and beyond*. Eastleigh, UK: John Libbey Publishing.
GCS (Government Communication Service). (2016, January). GCS evaluation framework. London, UK: Cabinet Office, Her Majesty's Government. Retrieved from https://gcs.civilservice.gov.uk/wp-content/uploads/2015/11/GCS_GCS-Evaluation-framework_A4-_191115.pdf
GCS (Government Communication Service). (2018, June). *GCS evaluation framework 2.0*. London, UK: Cabinet Office, Her Majesty's Government.
Gregory, A., & Watson, T. (2008). Defining the gap between research and practice in public relations programme evaluation: Towards a new research agenda. *Journal of Marketing Communications*, *24*(5), 337–350.
Gregory, A., & White, T. (2008). Introducing the Chartered Institute of Public Relations initiative. In B. van Ruler, A. Verčič, A., & D. Verčič (Eds.), *Public relations metrics: Research and evaluation* (pp. 307–317). New York, NY: Routledge.
Grunig, J. (2008). Conceptualizing quantitative research in public relations. In B. van Ruler, A. Verčič, & D. Verčič (Eds.), *Public relations metrics, research and evaluation* (pp. 88–119). New York, NY: Routledge.
Grunig, L., Grunig, J., & Dozier, D. (2002). *Excellent organizations and effective organizations: A study of communication management in three countries*. Mahwah, NJ: Lawrence Erlbaum.
Hatry, H., Houten, T., Plantz, M., & Greenway, M. (1996). *Measuring program outcomes: A practical approach*. Alexandria, VA: United Way of America.
Henningsen, A., Traverse Healy, K., Gregory, A. Johannsen, A., Allison, R., Bozeat, N., & Domaradzki, L. (2014, August 1). *Measuring the European Commission's communication: Technical and methodological report*. Retrieved from http://ec.europa.eu/dgs/communication/about/evaluation/index_en.htm
IAB [Interactive Advertising Bureau] (2009) Audience reach measurement guidelines. New York, NY. Available at http://www.iab.net/guidelines/508676/guidelines/audiencemeasurement
IPR (Institute for Public Relations). (2016). IPR Measurement Commission. Retrieved from http://www.instituteforpr.org/ipr-measurement-commission
Michaelson, D., & Stacks, D. (2011). Standardization in public relations measurement and evaluation. *Public Relations Journal*, *5*(2), 1–22.

Jenkins, N. (2015, September 28). Crowds gather to remember Hong Kong's umbrella revolution, one year on. *Time*. Retrieved from http://time.com/4051543/hong-kong-occupy-central-anniversary

Julian, D. (1997). The utilization of the logic model as a system level planning and evaluation device. *Evaluation and Program Planning, 20*(3), 251–257.

Kellogg Foundation. (2004). *Logic model development guide*. Battle Creek, MI: Author. Retrieved from http://www.smartgivers.org/uploads/logicmodelguidepdf.pdf (Original work published 1998)

Kellogg Foundation. (2010). *W. K. Kellogg Foundation evaluation handbook*. Battle Creek, MI: Author. Retrieved from https://www.wkkf.org/resource-directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook

Knowlton, L., & Phillips, C. (2013). *The logic models guidebook: Better strategies for great results* (2nd ed.). Thousand Oaks, CA: Sage.

Likely, F., & Watson, T. (2013). Measuring the edifice: Public relations measurement and evaluation practice over the course of 40 years. In K. Sriramesh, A. Zerfass, & J. Kim, (Eds.), *Public relations and communication management: Current trends and emerging topics* (pp. 143–162). New York, NY: Routledge.

Lindenmann, W. (2001). Research doesn't have to put you in the poorhouse. Gainesville, FL: Institute for Public Relations. Retrieved from http://www.instituteforpr.org/topics/research-savings

Lindenmann, W. (2003). Guidelines for measuring the effectiveness of PR programs and activities. Gainesville, FL: Institute for Public Relations. Retrieved from http://www.instituteforpr.org/wp-content/uploads/2002_MeasuringPrograms.pdf. (Revised version of original work publishd 1997)

Macey, W., & Schneider, B. (2008). The meaning of employee engagement. *Industrial and Organizational Psychology*, *1*(1), 3–30.

Macnamara, J. (1992). Evaluation of public relations: The Achilles Heel of the public relations profession. *International Public Relations Review*, *15*(2), 19–25.

Macnamara, J. (1999). Research in public relations: A review of the use of evaluation and formative research. *Asia Pacific Public Relations Journal*, *1*(2), 107–133.

Macnamara, J. (2000). The ad value of PR. *Asia Pacific Public Relations Journal*, *2*(1), 90–103.

Macnamara, J. (2002). Research and evaluation. In C. Tymson & P. Lazar, *The new Australian and New Zealand public relations manual* (21st Century ed., pp. 100–134). Sydney, NSW: Tymson Communications.

Macnamara, J. (2005). *Jim Macnamara's public relations handbook* (5th ed.). Sydney, NSW: Archipelago Press.

Macnamara, J. (2012). *Public relations theories, practices, critiques*. Sydney, NSW: Pearson.

Macnamara, J. (2014, June). Breaking the PR measurement and evaluation deadlock: A new approach and model. Paper presented to the Association for Measurement and Evaluation of Communication Summit on Measurement, Amsterdam. Retrieved from http://amecorg.com/downloads/amsterdam2014/Breaking-the-PR-Measurement-Deadlock-A-New-Approach-and-Model-Jim-Macnamara.pdf

Macnamara, J. (2015). Overcoming the measurement and evaluation deadlock: A new approach and model. *Journal of Communication Management*, *19*(4), 371–387.

Macnamara, J. (2016). A taxonomy of evaluation: Towards standards [Web site]. London, UK: Association for Measurement and Evaluation of Communication. Retrieved from http://amecorg.com/amecframework/taxonomy.html

Macnamara, J. (2018a). *Evaluating public communication: Exploring new models, standards, and best practice*. Abingdon, UK: Routledge.

Macnamara, J. (2018b). A review of new evaluation models for strategic communication: Progress and gaps. *International Journal of Strategic Communication*, 12(2), 180–195.

Macnamara, J., & Likely, F. (2017). Revisiting the disciplinary home of evaluation: New perspectives to inform PR evaluation standards. *Research Journal of the Institute for Public Relations*, *3*(2), 1–21. Retrieved from http://www.instituteforpr.org/research-journal

Macnamara, J., Lwin, M., Adi, A., & Zerfass, A. (2015). *Asia Pacific Communication Monitor 2015/16*: *The state of strategic communication and public relations in a region of rapid growth. Survey results from 23 countries*. Asia Pacific Association of Communication Directors (APACD) and Helios Media. Retrieved from http://www.communicationmonitor.asia

Marklein, T., & Paine, K. (2012, June). The march to standards. Presentation to the 4th European Summit on Measurement, Dublin, Ireland. Retrieved from http://amecorg.com/downloads/dublin2012/The-March-to-Social-Standards-Tim-Marklein-and-Katie-Paine.pdf

Marteau T., Ogilvie, D., Roland, M., Suhrcke, M., & Kelly, M. (2011). Judging nudging: Can nudging improve population health? *British Medical Journal, 342*(d228), 263–265.

McGuire, W. (1985). Attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology*, Vol. 2 (3rd ed., pp. 233–346). New York: Random House.

McGuire, W. (2001). Input and output variables currently promising for constructing persuasive communications. In R. E. Rice & C. K. Atkin (Eds.), *Public communication campaigns* (3rd ed., pp. 22–48). Thousand Oaks, CA: Sage.

McLaughlin, J., & Jordan, J. (1999). Logic models: a tool for telling your program's performance story. *Evaluation and Program Planning, 22*(1), 65–72.

Michaelson, D., & Stacks, D. (2011). Standardization in public relations measurement and evaluation. *Public Relations Journal*, *5*(2), 1–22.

Nesterak, E. (2014, July 13). Head of White House 'nudge unit' Maya Shankar speaks about newly formed social and behavioural sciences team. *ThePsychReport*. Retrieved from http://thepsychreport.com/current-events/head-of-white-house-nudge-unit-maya-shankar-speaks-about-newly-formed-us-social-and-behavioral-sciences-team

Pawson, R., & Tilley, N. (2001). Realistic evaluation bloodlines. *American Journal of Evaluation*, *22*, 317–324.

PRIA (Public Relations Institute of Australia). (2017). Measurement and evaluation framework. Sydney, NSW: Author. Retrieved from https://www.pria.com.au/resources/measurement-evaluation

PRSA (Public Relations Society of America). (2014). Measurement resources. Retrieved from http://www.prsa.org/intelligence/businesscase/measurementresources

Ragan/NASDAQ OMX Corporate Solutions. (2013). PR Measurement. Chicago, IL: Author. Retrieved from http://bit.ly/PRMeasureWPPR

Rhoades, L., Eisenberger, R., & Armeli, S. (2001). Affective commitment to the organization: The contribution of perceived organizational support. *Journal of Applied Psychology*, *86*, 825–836.

Rogers, P., Petrosino, A., Huebner; T., & Hacsi, T. (2000). Program theory evaluation: Practice, promise, and problems. In P. Rogers, A. Petrosino, T. Hacsi, & T. Huebner (Eds.), *Program theory evaluation: Challenges and opportunities* (pp. 5–13), New Directions for Evaluation, Vol. 87. San Francisco, CA: Jossey-Bass.

Rossi, P., Lipsey, M., & Freeman, H. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.

Schudson, M. (2003). Click here for democracy: A history and critique of an information-based model of citizenship. In H. Jenkins & D. Thorburn (Eds.), *Democracy and new media* (pp. 49–60). Cambridge, MA: MIT Press.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Stacks, D., & Bowen, S. (Eds.) (2013). *Dictionary of public relations measurement and research*. Gainesville FL: Institute for Public Relations. Retrieved from http://www.instituteforpr.org/topics/dictionary-of-public-relations-measurement-and-research

Sunstein, C., & Thaler, R. (2003). Libertarian paternalism is not an oxymoron. *University of Chicago Law Review, 70*(4), 1159–1202.

Taylor, M., & Kent, M. (2014). Dialogic engagement: Clarifying foundational concepts. *Journal of Public Relations Research, 26*(5), 384–398.

Taylor-Power, E., & Henert, E. (2008). *Developing a logic model: Teaching and training guide*. Madison, WI: University of Wisconsin. Retrieved from http://www.uwex.edu/ces/pdande/evaluation/pdf/lmguidecomplete.pdf

Thaler, R., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.

Valente, T. (2001). Evaluating communication campaigns. In R. Rice & C. Atkin (Eds.), *Public communication campaigns* (3rd ed., pp. 105–124). Thousand Oaks, CA: Sage.

Valente, T., & Kwan, P. (2013). Evaluating communication campaigns. In R. Rice & C. Atkin, *Public communication campaigns* (4th ed., pp. 83–97). Thousand Oaks, CA: Sage.

Watson, T., & Noble, P. (2014). *Evaluating public relations: A best practice guide to public relations planning, research and evaluation* (3rd ed.). London, UK: Kogan Page.

Weiner, M., & Bartholomew, D. (2006). Dispelling the myth of PR multipliers and other inflationary audience measures. Gainesville, FL: Institute for Public Relations. Retrieved from http://www.instituteforpr.org/wp-content/uploads/Dispelling_Myth_of_PR_Multiplier.pdf

Weiss, C. (1972). *Evaluation research: Methods of assessing program effectiveness*. Englewood Cliffs, NJ: Prentice Hall.

Wholey, J. (1987). Evaluability assessment: Developing program theory. In L. Bickman (Ed.), *Using program theory in evaluation* (pp. 77–92), New Directions for Evaluation, Vol. 33. San Francisco, CA: Jossey-Bass.

Wholey, J., Hatry, H., & Newcomer, K. (Eds.). (2010). *Handbook of practical program evaluation* (3rd ed.). San Francisco, CA: Jossey-Bass.

Wright, D., Gaunt, R., Leggetter, B., Daniels, M., & Zerfass, A. (2009). Global survey of communications measurement 2009 – final report. London, UK: Association for Measurement and Evaluation of Communication. Retrieved from http://amecorg.com/wp-content/uploads/2011/08/Global-Survey-Communications_Measurement-20091.pdf

Wright, D., & Hinson, M. (2012). Examining how social and emerging media have been used in public relations between 2006 and 2012: A longitudinal analysis. *Public Relations Journal, 6*(4), 1–40.

Zerfass, A., Verčič, D., Verhoeven, P., Moreno, A., & Tench, R. (2015). *European communication monitor 2015*. Brussels, Belgium: European Association for Communication Directors (EACD) and European Public Relations Education and Research Association (EUPRERA) in association with Helios Media, Berlin. Retrieved from http://www.zerfass.de/ECM-WEBSITE/media/ECM2015-Results-ChartVersion.pdf

Jim Macnamara PhD, FAMEC, FAMI, CPM, FPRIA is Professor of Public Communication at the University of Technology Sydney (UTS), a position he took up in 2007 following a distinguished career in professional practice spanning journalism, public relations, and media and communication research. Immediately before joining UTS, he was founder and CEO of the Asia Pacific office of the global media analysis firm, CARMA International for 11 years. He is the author of hundreds of articles and16 books including *Evaluating Public Communication: Exploring New Models, Standards, and Best Practice* published by Routledge in 2017. He is a Fellow of the International Association for Measurement and Evaluation of Communication (AMEC) and Chair of the AMEC Academic Advisory Group; a member of the Task Force on Standardization of Communication Planning and Evaluation Models; he developed the New South Wales Government Evaluation Framework for advertising and communication introduced in 2016; and he served as a member of the Evaluation Council of the UK Government Communication Service in 2016.

[1] The Global Alliance for Public Relations and Communication Management is a confederation of professional communication organizations representing 160,000 practitioners and academics worldwide (http://www.globalalliancepr.org).

[2] Causation is very difficult to establish in many cases, particularly when multiple influences contribute to *impact* (results), as is often the case. The three key rules of causation must be applied: (a) the alleged cause must precede the alleged effect/impact; (b) there must be a clear relationship between the alleged cause and effect (e.g., there must be evidence that the audience accessed and used information you provided); and (c) other possible causes must be ruled out as far as possible.

[3] Some include planning in *inputs*. However, if this occurs, formative research (which should precede planning) also needs to be included in *inputs*. However, most program evaluation models identify formative research and planning as key *activities* to be undertaken as part of the communication program. Inputs are generally pre-campaign/program.

[4] Reception refers to what information or messages are received by target audiences and is slightly different to exposure. For example, an audience might be exposed to a story in media that they access, but skip over the story and not receive the information. Similarly, they may attend an event such as a trade show and be exposed to content, but not receive information or messages (e.g., through inattention or selection of content to focus on).

[5] Learning (i.e., acquisition of knowledge) is not required in all cases. However, in some public communication campaigns and projects it is. For example, health campaigns to promote calcium-rich food and supplements to reduce osteoporosis among women found that, first, women had to be 'educated' about osteoporosis (what it is, its causes, etc.). Similarly, combatting obesity requires dietary education. Whereas *understanding* refers to comprehension of messages communicated, learning refers to the acquisition of deeper or broader knowledge that is necessary to achieve the objectives.

[6] OTS is an abbreviation of 'opportunities to see', usually calculated the same as impressions or gross audience reach.

[7] Net promoter score is a score out of 10 based to a single question: 'How likely is it that you would recommend [brand] to a friend or colleague?' Scores of 0–6 are considered 'detractors'/dissatisfied; scores of 7–8 are satisfied but unenthusiastic; and scores of 9–10 are those considered loyal enthusiasts, supporters, and advocates. (See https://www.netpromoter.com/know)

[8] Econometrics is the application of mathematics and statistical methods to test hypotheses and identify the economic relations between factors based on empirical data.