Funded by the European Union

EUROMED MIGRATION
هجرة

Implemented by
ICMPD
International Centre for Migration Policy Development

How to Perform Impact Assessments:

Key Steps for Assessing Communication Interventions

Developed for ICMPD under the EU funded programme EUROMED Migration IV by the Observatory of Public Attitudes to Migration - Migration Policy Centre, European University Institute, Florence

Dr. James Dennison

http://migrationpolicycentre.eu/opam/about
http://migrationpolicycentre.eu

The presented study includes active links to sources and references that are accessible online.

Readers are strongly encouraged to consult the soft version of this study to access all proper links.

Design: Pietro Bruni - www.toshi.ltd

Executive Summary

- The demand for assessing the impact of communication campaigns across all sectors is increasing. In the field of migration, this demand is particularly acute because of, amongst other things, the potential of such campaigns to have real-world consequences for millions of people.

- Despite high demand, there is limited understanding amongst communication practitioners about how to perform impact assessments (IAs), in both the private and public sectors, partly due to an historic lack of consensus among experts.

- Migration professionals have both reported significant difficulties in performing such assessments—not least because of the, often unrealistic, methodological standards of academia or international organisations—and had their attempts at such assessments severely criticised.

- This report brings together disparate terminology, findings and recommendations from the private and public sectors and academia to synthesise a set of five general steps for practitioners when performing IAs.

- IA compares a variable of interest after an intervention (e.g. a communication campaign) to what it would it be if that intervention had never happened.

- Accompanying each of the five steps are examples of how migration communicators can assess the impact of campaigns to, first, change attitudes and, second, change behaviour.

- Step 1: Set objectives for the intervention that define what the desired effect is. Ideally, this should (1) meet the SMART criteria of specific, measurable, attainable, relevant and time-bound; (2) focus on ultimate outcomes, rather than only outputs; (3) avoid a number of common mistakes in setting objectives and (4) carefully consider what the—as specific as possible—target audience is.

- Step 2: Identify a measure of the impact(s) that the intervention has as its objective. These should be (1) clearly defined and unambiguous; (2) externally valid (they act reliably if used for other interventions) and (3) internally valid (they measure what they claim to measure). There already exist long-lists of measures (also known as indicators) to choose from.

- Step 3: Estimate what would have happened had there been no intervention, known as a 'counterfactual'. Ideally, but not necessarily, this requires performing a pre-intervention mea-

surement, also known as baseline assessment. However, this is not always possible. This should not dissuade practitioners but instead assumptions of any IA, as discussed below, should be made openly. A long list of methods for measurement, including sampling considerations, are provided.

- Step 4: Perform the intervention. This may involve separating a randomised sample into a treatment group(s)—that receives the intervention(s)—and a control group—that does not. If this is not feasible, other forms of counterfactual are possible that 'construct' a control group. This section also overviews: (1) types of communication campaigns; (2) the MINDSPACE checklist of behavioural influences for interventions; (3) key recommendations on persuasive interventions from ICMPD's previous work; (4) recommendations on campaigns deterring irregular migration.

- Step 5: Post-intervention measurement and analysis: (1) those receiving the intervention should be measured post-intervention (ideally, as well as, either a genuine or constructed control group); (2) the impact must be calculated: usually the difference in the variable of interest pre- and post-intervention and/or between the treatment and control group; (3) theoretical consideration of why and what aspect of the campaign caused an impact; are the findings likely to be universalizable?; (4) creation (and, in some cases, incorporation) of recommendations based on combination of impact and theoretical considerations.

Conceptual and theoretical starting points

- IA is concerned with asking: How are things different compared to if there had been no intervention?
- What 'things' should we consider?:
    - What is the objective of the IA? (Step 1).
    - How can we measure these 'things'? (Step 2)
- How can we know 'how things would be if there had been no intervention'? (Step 3 and 4)
    - First, how were things before the intervention?
    - But (!) something else besides the intervention may have happened in the meantime.
    - Second, how are things different in an identical (control) group that has not received an intervention?
    - But (!) humans (and their experiences) are all unique so no truly identical group can exist.
    - Best to do both: pre- and post-measurement of both a 'control' and a 'treatment' group!
    - But (!) this is expensive and not always possible. If neither a pre-intervention measurement nor a control group can be measured, a 'control group' can be constructed that is as close as possible in theoretically important ways to the treatment group! This group can be then measured.
- How are things after the intervention? (Step 5)
    - This one is relatively easy – measure the 'things' of interest amongst those (or a representative sample of those) who have received the treatment. Then consider the how the pre- and post-intervention measurements are different for both control and treatment groups. This difference is the impact!

- Actually, we can never know with certainty 'how things would be if there had been no intervention'! As such, we are not looking for absolute truths but rather the least bad estimation possible. There is no one way to do this, but it should at least be based on some systematic method from which it would be possible to conclude that no impact has occurred. Furthermore, we should not be ashamed to be open about our assumptions and our IA's potential shortcomings. In fact, we should be more suspicious of IAs in which none of these are explicitly stated. This, rather than 'absolute truths', is good practice and genuine science.

- Why did the intervention have this effect? Answering this is necessary to understand in what other circumstances the intervention is likely to have the same effect and to, thereafter, make recommendations for future interventions! This consideration requires theory. Luckily, theory is essentially common sense. However, even more so than regarding how things would be different otherwise, we can be even less sure about why things would be different otherwise—as such we should be open about our theoretical and methodological assumptions, caveats and shortcomings.

Table of contents

Introduction and theoretical considerations

This report considers how to perform Impact Assessments (IAs), with a particular focus on communication campaigns, projects and activities (collectively referred to as interventions) in the public policy field of international migration. IAs–the formal estimation of the effect of a policy intervention—of communication campaigns are both of vital practical importance and of acute scientific interest. Communication campaigns have real-world consequences for millions of people. The scale and economic importance of migration in the twenty-first century makes the strategic decisions faced by communicators on this subject of considerable financial impact. More profoundly, the legal and rights-based imperatives when governing migration underscore the humanitarian and political importance of communicating in the most effective way possible on the subject. This same political importance of communicating well is granted yet further gravity, and complexity, by the often highly charged political questions of collective identity, values and community that discussing the topic of migration engenders.

Given that the aims of communicating on issues of migration are so often related to influencing public attitudes, such as correcting perceptions or affecting favourability to immigrants, or behaviours, such as propensity to migrate, IAs can also produce deeper insights into key social scientific questions, helping to explain why humans do as they do and think as they think. They can provide evidence of what motivates humans and how, for example, trust or willingness to take risks are formed. Their outcomes may also provide evidence in support of, at times, competing scientific theories that see public opinion as, on the one hand, volatile, irrational and prone to manipulation or, on the other, the result of deep-seated and stubborn psychological predispositions or early-life experiences. More sophisticated, contemporary theories that consider the interactions between these forces can also be tested by proper impact assessment, the conduction of which—usually the result of work in the real-world—also raises theoretically profound methodological questions of identifying causality.

Unsurprisingly, then, as governments and advocacy organisations have, in recent decades, poured additional resources into communications campaigns, they are now demanding that their public sector communications campaigns are shown to be 'necessary, cost-efficient and effective' (Macnamara, 2020a: 361). However, there remains a 'stasis' in understanding how to measure and evaluate such campaigns (Gregory and Watson, 2008). This lack of understanding is not restricted to migration communication nor even the public

sector. Scholars have lamented the lack of proper impact assessment in the fields of health communication and health promotion (Noblet and LaMontagne, 2009). In the public relations industry, there remains no consensus on 'what the basic evaluative measures are or how to conduct the underlying research for evaluating and measuring public relations performance' (Michaelson and Stacks, 2011: 1). Strikingly, despite its $500billion annual expenditure, 'The advertising industry, as a whole, has the poorest quality assurance systems and turns out the most inconsistent product … of any industry in the world' (Thomas, 2008: 1; cited in Macnamara, 2020a). As such, effective impact assessment, a component of (ex-post) evaluation along with the likes of the better understood cost-benefit analyses, remains the 'holy grail' for public relations (Gregory, 2020).

The dearth of effective evaluation of communications campaigns is particularly apparent in migration campaigns. For the sake of this report, we consider such campaigns that roughly split into two key categories: those informing would-be irregular migrants about the risks of such a journey (often with the aim of decreasing irregular immigration) and those attempting to inform or change the attitudes of host populations to regular migration and immigrants (usually with the aim, implicit or explicit, of decreasing negativity or "polarisation"). Such campaigns have received significant funding in recent years, both due to the 'assumption that a lack of accurate information generates irrational and risky irregular migration behaviours' (Browne, 2015: 1) and the high public salience of immigration as a political issue in western democracies (see Dennison and Geddes, 2019; Dennison, 2019), respectively. However, 'there is extremely little evidence on the impact and effectiveness of these campaigns' with very few evaluations being made publicly available and those that are being typically of fairly low reliability (Brown, 2015: 2). Moreover, identifying a causal effect of a specific campaign on such objectives has been described as 'difficult' or 'almost impossible' by the European Migration Network (2012). Meanwhile, academic and scientific recommendations for such impact assessments have often severely underestimated the financial and logistical difficulties in performing them.

Indeed, a number of reasons for the lack (and lack of quality) of impact assessment in communication campaigns have been offered, including a lack of (1) budget, (2) demand, (3) standards and (4) knowledge (see Macnamara, 2020: 362, for list of sources). However, with an increasing tendency for government communications spending and grants to demand impact assessments (e.g. since 2014, all work by the UK Government Communication service) and with the previously 'raging' academic debate over how to measure and evaluate such campaigns now beginning to reach some level of consensus (Gregory, 2020: 367), the only explanation that is still plausible is a lack of knowledge. Increasing such knowledge is thus the aim of this report.

To do so, the report draws on findings from the academic and public policy literatures, both from the field of migration (e.g. Tjaden et al, 2018) and elsewhere (e.g. WHO, 2017; European Commission, 2017; UK Government Communications Service, 2018) as well as findings drawn from the private sector (e.g. AMEC, 2020), to produce a guide for the practices and considerations needed to produce reliable impact assessments of communication campaigns. The result is a five-step plan.

Readers may notice that, rather than IAs being something done after an intervention or activity, for IA to work, the most important steps must actually be taken prior to the intervention. Though this plan is by no means exhaustive it does bring together the major findings from the sources listed above as well as broader social scientific considerations. As such, performing such an impact assessment should result in reliable evidence and conclusions regarding the effect of a specific intervention. However, as mentioned, public policy professionals are not always able to meet the very highest standards of robustness as outlined by academic research and the guidelines of large international organisations, such as the World Bank, due to financial reasons or the logistical realities of complex real-world action. As such, the report also demonstrates the implications of sometimes necessary deviations from so-called 'gold standard' of measuring such changes. Overall, the report proceeds with five sections that elaborate on each of the five steps outlined above.

Before that, however, we must briefly overview the conceptual and theoretical considerations of performing IAs. In short, this report contends that Impact Assessment is essentially concerned with asking 'how are things different after an intervention compared to if there had been no intervention?'. The hypothetical nature of this question already hints at the necessarily imperfect nature of any IA, regardless of its seeming complexity of robustness. From this question, we can also derive the main steps and challenges of performing IAs as expressed through a number of sub-questions. First, what 'things' do we mean? Section 1 argues that choosing these should be determined solely by the objectives of the IA while Section 2 considers how to then measure these things. Second, how can we know 'how things would be if there had been no intervention'? There are two primary ways to do this. The first is to consider how things were prior to the intervention; however, this might then capture the effect of something else that happened during the intervention, resulting in a spurious finding. The second is to consider how things are in an identical "control group" that did not receive the intervention. However, amongst humans no *truly* identical control group can be created. Therefore, the ideal, or "gold standard", is to do both, i.e. measure *both* before and after in *both* a "treatment group" and a "control group". However, this is expensive and not always possible. If *neither* a pre-intervention measurement *nor* a control group can be measured, a "control group" can be constructed that is as close as possible in theoretically important ways to the treatment group. This group can be then measured. These issues are covered in sections 3 and 4. The third sub-question is to consider how things are after the intervention. This is relatively simpler and covered in section 5.

Two important theoretical points are worth raising at this stage regarding both description ("how things are different?") and, crucially for learning lessons for future interventions ("why things are different?"). First, as mentioned, we can never know with certainty 'how things would be if there had been no intervention'. As such, we are not looking for absolute truths but rather the least bad estimation possible. There is no one way to do this, but it should at least be based on some systematic method from which it would be possible to conclude that no impact has occurred (in social science this is known as a falsifiable hypothesis). Furthermore, we should not be ashamed to be open about our assumptions and our IA's potential shortcomings. In fact, we should be more suspicious of IAs in which none of these are explicitly stated. This, rather than 'absolute truths', is good practice and genuine science. Second, attempting to answer why the intervention had the observed impact is necessary to understand in what other circumstances the intervention is likely to have the same effect and to, thereafter, make recommendations for future interventions. This consideration requires theory. Luckily, theory is, to an extent, often common sense. However, even more so than regarding how things would be different otherwise, we can be even less sure about *why* things would be different otherwise—as such we should be open about our theoretical *and* methodological assumptions, caveats, shortcomings, etc.

1. Specify objectives

Academics and practitioners are essentially united in the necessity of setting clear objectives for any communication campaign before any impact assessment, or broader evaluation, can be possible. Gregory (2020: 369) outlines the five steps of the 'communication planning cycle', that can be applied to business planning more generally, starting with the imperative to 'define aims or goals'. Similarly, the UK Government Communications Service's (GCS) model of strategic planning, known by the acronym of OASIS, begins with setting 'Objectives'. According to Gregory (2020), setting objectives, alongside setting a benchmark (see steps 2 and 3 of this guide) form part of the broader pre-intervention third of any evaluation (which IA is one type of), known as the 'formative' stage. From, primarily, the private sector, the Association for Measurement and Evaluation of Communication (AMEC, 2020) have also produced their own seven 'Barcelona Principles' of impact assessment. Instead of being a step-by-step guide, these principles should be kept in mind throughout the process of evaluation. A number of these relate to setting objectives, most importantly: 'setting measurable goals is an absolute prerequisite to communication planning, measurement and evaluation'. The reason for the absolute importance of identifying objectives is simple: all form of measurement and analysis derive from knowing what type of impact we are interested in assessing. This type of impact should be in accordance with the ultimate objectives of the intervention.

The nature of objective setting is typically differentiated between organisational objectives—which are broader and more long-term—and 'single communication activities' which we are primarily concerned with regarding IAs (European Commission, 2017). Regarding the latter, all of Gregory (2020), the European Commission (2017: 32), GCS (2018) and AMEC (2020) argue that such objective setting should meet the, by now well-known, SMART criteria: specific, measurable, attainable, relevant and time-bound. Henningson et al (2014: 1), writing for the European Commission, describe this as 'a baseline for best practice communication measurement and evaluation'. The UK's GCS add an additional 'C' to the SMART criteria, standing for 'challenging'.

In line with this logic, the World Health Organisation (WHO) (2017: 42) encourages communicators to 'start evaluation with realistic objectives', noting that immediate behavioural changes are unlikely but 'communicators can show incremental improvements … [related to] … increasing awareness, enhancing understanding

of health risks, or nurturing skills.' Similarly following SMART criteria, the European Commission (2017: 29) argues that communicators should 'be as specific as you possibly can' in defining who the target audience is, the desired behavioural change, their current attitudes (to your policy) and who their influencers are.' Moreover, they (2017: 86) point to a number of key questions to ask when objective setting on single communication activities:

- What is the EC [or relevant organisation] trying to achieve (i.e. 'the big picture')?

- 'What do you want to achieve by this action, specifically?' 'What is the measurable purpose of this action? If there is no real and specific purpose, then there is a strong chance that the action will be: a) of no real benefit to the citizens of Europe; and b) could be wasting taxpayer's money.'

- 'Who are you trying to reach?' '"The public" is not an acceptable target', however, '16-18 year olds who could be participants in an academic exchange programme' is.

Both the WHO (2017: 45) and European Commission (2017: 28-29) list a number of examples of objectives of communication changes; respectively, 'raising awareness, increasing knowledge, influencing attitudes, and building confidence in WHO recommended changes' and helping policy by persuading, informing, normalising, inspiring and engaging.

However, the latter warm against 'awareness-raising' as an objective, which it often is in migration communication, as being overly vague. Similarly, though now explicitly specific to migration communication, Trajden (2018: 7) note that 'a common issue is the lack of a clearly defined campaign objective and/or target group. This hampers any rigorous evaluation of programme effects. Whenever an objective is defined, it is most often aimed at "awareness-raising" and "knowledge generation", neither of which are then properly defined.' Awareness-raising is a particularly problematic objective, first, because it is rarely linked to broader organisation objectives, which usually, in the case of migration, relate to changing attitudes and behaviour and, second, because the campaign can then be evaluated more in terms of outputs (e.g. some amount of advertising or publicity) rather than outcomes.

Specifically, regarding the link between single communication activities and broader organisational objectives, Zerfass and Volk (2020) lament the lack of any systematic system or criteria by which the link between a single communication activity and broader goals can be judged.

Finally, regarding key contemporary shifts in good practice, Macnamara (2020b) lists ten ongoing problems in the impact assessment of communication campaigns. More than any other aspect of impact assessment, these relate to the objective setting stage. These are:

- A focus on outputs (what was done?) rather than outcomes (how did things change?) or, better still, impact ((how) did the communication campaign contribute to the change?).

- Related, assumptions about effects. In short, this is assuming that outcomes and impact are the same—i.e. that any desired real-world changes were the result of the campaign (see also, confirmation bias). It is for this reason that the measurement of a control group forms part of best practice.

- Related again, a focus on measurement rather than evaluation. The former is essentially simply data-collection, whereas making judgements about 'impact' and whether objectives were met requires judgement and theoretical considerations (see step 5).

- Seeing evaluation as something done purely after the campaign. Amongst other things, this leads to post-facto objective setting and justifications and a lack of appropriate pre-intervention measurements.

- Invalid and spurious methods.

- Media-centricity. This is the tendency for those working in communications to focus on what they know when setting objectives, i.e. the world of media. This results in outputs disproportionally focussed on media engagement, whether in traditional or social media, such as mentions, shares, viewers, etc., rather than actual impact judged according to objectives that contribute to organisational goals.

- Disciplinary, sectoral or industrial siloes. The tendency for communication evaluation to only follow guidance from those doing similar work, e.g. migration practitioners only working on that.

- Lack of knowledge and skills. The fact that most individuals working in communication know how to communicate but not how to link it to broader strategic goals, reducing the impetus for evaluation.

- Quantitative bias. An over-focus on easily-accessible numerical data, rather than more nuanced, qualitative data and evidence.

2. Identify measure of impact

The second step of communications impact assessment is identifying an appropriate measure of impact, as inferred from the campaign's objectives. In addition, some impact assessment requires communications to determine what success would look like, in the terms of the identified measure of impact (for example, a one per cent increase in the proportion of individuals correctly responding to a survey on migration figures). The European Commission's "Toolkit for the evaluation of communication activities" (2017: 18) defines those measures, or indicators, as 'a quantitative or qualitative factor or variable that provides a reliable measurement which reflects the change of an intervention. They are specific observable and measurable characteristics. An indicator is about "how much" or "how many" or "to what extent" or "what size".' AMEC's (2020) second 'Barcelona Principle' states that 'Measurement and Evaluation Should Identify Outputs, Outcomes, and Potential Impact.' Thereafter, they divide measurement into planning and execution. For the measurement planning stage, their most important recommendations are: (1) to 'identify concrete (data) sources tailored to align with each of the' objectives; (2) prior to any intervention, all measurement 'tools and tracking methods, KPIs and benchmarks' should be identified; (3) identify both outputs ('metrics/products that result from communication') and outcomes ('changes in beliefs, attitudes, behaviours'). In terms of measurement execution, they recommend 'aim for measuring outcomes. However, outputs do play an important role in helping to understand the holistic picture', while they also recommend to link both to the broader context and keep the evaluation and measurement ongoing.

In terms of choosing indicators, the following seven steps align with the recommendations of the European Commission (2017) and other organisations while being largely feasible, even for relatively small organisations, even if they are more focussed on broader, multi-activity evaluation than the single activity impact assessment we are concerned with here:

1. Choose indicators prior to any activity
2. Select respective indicators for each of the five levels of activity, defined (somewhat confusingly) as:
    a. Outcome: did you achieve your overall objectives
    b. Intermediary outcome: what did the target audience do as a result
    c. Outtakes: immediate effect on target (learning, awareness, etc.)
    d. Outputs: the people you expect to reach directly with your activity
    e. Inputs: the activities carried out.
3. Each communication tool should have its respective indicator
4. The indicator should directly reflect the target audience
5. The indicator should ensure comparability across time (notably before and after the intervention) but also with relevant findings elsewhere if possible.
6. Devote sufficient resources to measurement
7. Prioritise indicators that are closest to objectives.

However, impact assessment will not always be able to be on the same scale as that required by the European Commission, making a large set of indicators not possible and, instead, the necessity to focus on one key measurement. The WHO (2017) specifies the desirable characteristics of indicators as:

- clearly defined and beyond interpretation;
- reliable (i.e. if the indicators are used again under separate but identical circumstances, they will produce the same result, known as external validity in social science);
- valid (internally; i.e. it measures what it claims to measure); quantifiable; financially and practicably feasible.

Gregory (2020: 374) offers a number of examples, with measurement methodologies for each of inputs, outputs and outcomes. While the former relate to both researching the intervention as well as the nature of the intervention itself, outputs relate to the measurements of who has received (or understood, retained, considered, been given, etc.) the intervention, while outcomes are almost always related to the number or proportion who changed their attitudes or behaviour. Whereas all of these form a part of evaluation, IA is primarily concerned with some form of outcome. The methodologies for these typically are one of survey data, "big" (i.e. incidental) data, focus groups, semi-structured interviews, content analysis, ethnographical observation or official data, records or secondary statistics.

In terms of specific indicators, both the academic literature and public policy documents have attempted to come up with various quite exhaustive lists. Romenti and Murtarelli (2020: 394) offer a number of communications indicators to measure social media outputs: 'Clickthrough rate (CTR); % Referral traffic; Coupon download (% increase or decrease rate); Information request (% increase or decrease rate); Really Simple Syndication (RSS) subscriber (% increase or decrease rate); Number of rating; % Sharing rate per post; % Retweet rate per post.'

O'Neil (2020: 407) shows how measures can be linked to objectives and methods in this example 'outcomes evaluation framework', as shown in Table 1.:

Table 1. Example 'outcomes evaluation framework'

Communication objectives	Outcomes (in terms of measurement)	Type of outcomes	Evaluation methods	Sampling frame
In 2016, the majority of citizens of province A are aware and supportive of issue X.	80% of citizens of province A are aware of issue X.	Awareness	Survey	Survey: Representative sample of citizens of province ABC
	Citizens express support for issue X	Attitude	Survey Focus groups	Survey: Representative sample of citizens of province ABC Focus groups: group discussions with selection of key demographics

The New South Wales government offers a similar five levels of activity—this time 'inputs, activities, outputs, outcomes and impact—with a concise list of things that may be measured for each and how these can be measured (DPC, 2016). Most pertinently, those of the latter three are shown in the table below:

Table 2. NSW's examples of items to be measured for outputs, outcomes and impact.

Level of activity	Item to be measured	Examples of measurement
Outputs	paid advertising, media publicity, publications, websites, social media posts, events, sponsorships, community projects	media metrics (reach, impressions, OTS, TARPs, etc.), content analysis, website data, social media statistics, reader surveys
Outcomes	recall, awareness, interest, engagement, preference, attitude change, satisfaction, trust, intentions, advocacy	Surveys, interviews, social media analysis, ethnography, net promoter score
Impact	behaviour, inquiries or registrations, revenue, retention, quality of life/well-being, savings	Behaviour tracking (e.g. databases), surveys, CBA/ROI, well-being metrics

Finally, the UK's GCS (2018) also lists things that may need to be measured for each of its 'inputs, outputs, outtakes, and outcomes'. These are:

Table 3. UK's GCS's (2018) examples of items to be measured for outputs, outcomes and impact.

Level of activity	Item to be measured	Examples of measurement
Input (what is done before and during the activity)	Planning	£, FTE days, Volume by type, no. of press releases sent out, no. of releases to social media channels
	Preparation	
	Pre-testing	
	Production	
Output (what is delivered)	Distribution	Aggregate audience (offline/online) reach; Volume of events; volume of press exposure; number of partnerships
	Exposure	
	Reach	
Outtakes (what target audience do with the output)	Awareness	% of impressions generating an interaction / reaction; CTR; VTR; average time spent on website; % of visitors that navigate; recall rate; proportion that indeed to act in accordance with campaign; etc. ; Sentiment towards campaign; change in contact / impressions
	Understanding	
	Interest	
	Engagement	
	Preference	
	Support	
Outcomes (result of the activity on the target audience)	Impact	Number / proportion that change behaviour; number of successful recruitments; change in expressions of interests; unit costs; change in attitudes
	Influence	
	Effect on Attitudes	
	Effect on Behaviour	
	Revenue	
	Cost reduction	
	Complying actions (attitude/behaviour change)	
	Retention	
	Reputation	

3. Counterfactual measurement

In order to assess the impact of any campaign, project or intervention, some measurement of what the object of change would have been if there had been no intervention is necessary. Ideally, though not necessarily, this should involve pre-intervention measurement. This is sometimes known as 'baseline assessment', which are 'carried out before a new tactic or intervention is used to improve performance. Ideally, the source and method of collection of this data should be as similar as possible to, if not exactly the same as, that of the post-intervention measurement. However, for various reasons this is also not always possible. Furthermore, at times pre-intervention measurement is combined with objective setting, so that the latter is inductive (the objectives are based on an assessment of the current empirical reality) rather than deductive (based on logic or theory).

Although impact assessment can be carried out using an endless number of potential research designs, when assessing the impact of communication campaigns, particularly in the field of migration, Tjaden et al (2018: 16) recommend randomised control (or controlled) trials (RCTs) as the 'gold standard'. In short, RCTs see a representative sample of the target population randomly allocated into a control group and one or more treatment (i.e. receiving an intervention) groups. Both/all groups are measured pre-intervention and post-intervention, with the difference between these in the treatment group representing the 'impact' of the intervention (presuming there is no change in the control group). The randomised treatment allocation ensures that there are no systematic (or at least theoretically important) differences between the control and treatment groups.

However, such a robust method is not always feasible for those working in practice. First, most real-world interventions—the most prescient subject of IAs—are unable to confidently know who has received the intervention or not, unlike the medical trials from which the lauded RCTs originate. For example, a public campaign cannot know with certainty who has received the treatment and who has not, let alone be sure that the receipt of such a treatment is random. As such, Tjaden et al (2018) rank methods whereby each method incrementally removes one of the assumptions of a genuine RCT, so that the highest rank (1) see randomised treatment allocation, control-group design, pre- and post- measurements and a large sample size (i.e. an RCT); (2) removes randomised treatment allocation; (3) also removes control-group design; (4) also removes pre- and post- measurements and (5) also removes large sample size.

As this suggests, pre-intervention measurement is not necessary for some form of impact assessment, although it is desirable. Without pre-intervention measurement, some other form of 'baseline' or counter-

factual must be created. This can be via 'quasi-experimental' designs, essentially via the construction of a control group through various advanced research methods (matching, regression discontinuity, propensity scores, etc.; see Gertler et al, 2016, for an explanation of these). Far more common, non-experimental designs look systematically at whether the evidence available suggests that there has been an impact and considers whether other factors could be causing the suggested impact. Pre-intervention measurement can also be replaced in these cases with other sources. In this case, the post-intervention measurement should be designed to be as similar as possible to these other sources, for example, using the same questions as international surveys, which is used as the pre-intervention measure.

As mentioned in the previous section, there are numerous potential methods for collecting the data that represents the pre-intervention measure: surveys, focus groups, interviews, experiments, ethnographic studies, 'observation studies', case studies, tracking mechanisms and web metrics (O'Neil, 2020: 408). The European Commission (2017: 59) offers details of a number of more specific forms of measurement that can be used in impact assessment:

- Qualitative research tools: case studies; focus groups (experts / stakeholders / direct target audiences / electronic); participatory observation; stakeholder interviews; diaries
- Online: social media monitoring; web analytics; website usability audit; web-visibility mapping
- Surveys: paper/print; online; face-to-face; telephone; tracking studies
- 'Other tools': content mapping/audit; customer [user] journey mapping; mystery shopping; organisational review; technical audit; data mining
- Secondary data: benchmarking; analysis of secondary data; analysis of quantitative data; analysis of qualitative data; expert panels; SWOT analysis; multi-criteria analysis
- Media analysis: quantitative; qualitative
- Advertising: recall; frequency; reach and coverage; OTS/OTH; ratings; cost; share
- Cost analysis: ROI; CBA; Cost effectiveness / efficiency analysis.

When considering the major impacts being assessed in migration research, we are usually concerned with changing attitudes or behaviour, as final outcomes. Many of the above are naturally related to outputs, particularly those monitoring media and advertising. However, there is still plenty of choice amongst potential forms of measurement for out two ultimate outcomes of interest. Some examples for each include:

- Attitudes to immigration: most commonly surveys. However, see also: focus groups; interviews;

social media monitoring; 'revealed preferences', etc.

- Propensity to migrate irregularly: most commonly official statistics. However, see also surveys; focus groups; interviews; social media monitoring, "Big data".

Finally, many of the above forms of data collection will raise the issue of sample sizes. In short, a sample is necessary whenever it is not feasible to measure the entire population of interest. The population of interest directly derives from the intervention's objectives. For example, if a communication campaign is interested in changing attitudes to immigration amongst 18-24 year olds in region X of country Y, it is unlikely that the campaign will have the resources to measure attitudes before and after for all 18-24 year olds in that region. Instead, we can find a representative sample of the population and infer from any changes observed in this group that such changes have / would have occurred in the general population. How can we attain all-important representativeness? Gartler et al (2016) and Tjaden et al (2018) suggest that the best way is to ensure a sufficiently large sample. Though increasing the size of the might improve representativeness, it is neither sufficient nor necessarily true. Instead, representativeness comes from being as similar to the general population as possible. However, such similarity can of course mean many things. In general, we want the sample to be as similar to the general population in terms of the various theoretically plausible predictors of whatever the outcome of interest is. So, if we are interested in measuring impact on attitudes to immigration, we should make sure that the sample resembles the general population as closely as possible in terms of all of those things (or as many as possible) that have been shown to affect attitudes to immigration. These would therefore include at least some of:

- Socio-demographics and socialisation: gender, age, education, income, region, ethnicity, immigration background, family
- Political profile: left-right self-placement, party identification
- Context: media consumption; local immigration rates

4. Perform the intervention

Although this report is concerned with impact assessment rather than advising on the specifics of what an intervention regarding migration should look like, the latter does raise a number of issues related to impact assessment. First, in order for experimental research to be carried out, as already discussed, any intervention should be isolated to a 'treatment group'. Ideally, again, membership of this group, rather than the 'control group' should be randomly allocated from a representative sample.

Trajden (2018: 12) lists potential communication tools or mediums that comprise such interventions, the impacts of which can then be assessed as: 'websites, social media, TV shows and video productions, radio reports, print media (including newspaper articles, posters, billboards, postcards and flyers), workshop-type activities, parades, concerts, (road)shows, quiz programmes, comic strips/ books, theatre, sports-related activities, hotlines and information centres, and word-of-mouth peer networks.

Macnamara (2020b: 449) recommends drawing on behavioural insights when designing campaigns, in particular the "MINDSPACE" checklist of influences on human behaviour. These "nudge-theory"-inspired influences are:

- Messenger: the influence of who communicates the information
- Incentives: the influence of mental shortcuts such as loss avoidance and present bias
- Norms: the influence of what others do.
- Defaults: the tendency to go with the flow
- Salience: the influence of what seems important and/or novel
- Priming: the influence of subconscious cues
- Affect: the influence of our emotional associations
- Commitments: our desire to be consistent with promises we have made and duties
- Ego: the influence of what makes us feel good about ourselves.

Specifically regarding interventions aimed at changing attitudes to immigration, Dennison (2020) collates the findings of a number reports, as shown below.

Table 4. Summary of key recommendations from existing best-practice guides for migration communication

	Sharif (2019)	Banulescu-Bogdan (2018)	Marthouz (2006)	Bamberg (2018)	Welcoming America (2018)	Christiano (2017)
Strategic arrangements						
Develop a proactive communications strategy	X	X				
Set up partnerships for communications/support others	X		X			
Research and target moveable audience, know their perceptions and prejudices	X		X	X	X	
Communications content						
Focus on values	X	X	X	X	X	X
Appeals to emotion	X				X	X
Hope/positivity/solutions focus	X		X			
Avoid attacking audience		X				
Avoid repeating opposing ideas / increasing their salience		X		X		
Find common ground		X				
Neutralise opposition arguments			X			
Use storytelling				X		X
Communications delivery						
Choose credible messengers, including migrants or moderates	X	X				
Use succinct / digestible / focussed messaging	X			X	X	X
Be visual	X					X

Finally, specifically regarding 'communication campaigns to deter irregular migration from developing to developed countries', Brown (2015: 1, 3) citing the UNHCR (2011), lists recommendations regarding how to perform such information campaigns:

- Information campaigns are most effective when they target the entire community rather than only potential migrants, since decisions to leave a home country are generally based on, and supported by, a family or community.

- Information campaigns must not discourage legitimate refugees or asylum seekers. Raising awareness about legal migration opportunities, where they exist, can increase the effectiveness of information campaigns.

- Mass media campaigns, using radio or television, can address large audiences of different profiles and backgrounds.

- Discussion sessions and theatre productions may reach fewer persons, but they offer a more in-depth opportunity to discuss, exchange ideas and persuade individuals to change their minds.

- "Catch-phrase messages" are useful for attracting the attention of the audience and providing information on complex matters in a direct and memorable manner. The language of these messages can also be tailored to the culture of the audience.

- Real-life testimonies can render information more accessible and intelligible.

- Using celebrities or high-profile individuals to convey messages can help establish trust, reach the target audience, and raise difficult and sometimes contentious issues.

Their main lessons for stakeholders are:

- Disseminate information, education and communication materials to inform individuals of the risks of irregular movements, including human trafficking and smuggling.

- Use simple targeted messages that are age, gender and culture sensitive and translated into appropriate languages to reach a wide audience.

- Encourage the involvement of persons who have experienced hazardous journeys to help influence individual choices and shift attitudes of host communities.

- Include contact details for support services in information leaflets, as well as the rights and obligations of persons on the move and available international protection and legal migration options.

- Initiate awareness-raising activities on the plight of refugees and the protection needs of persons travelling within mixed movements.

- Involve law enforcement, government officials, politicians and local communities in information strategies, and encourage open debate to identify outcomes for persons travelling within mixed movements.

- Cooperate with relevant actors on developing information campaigns in countries of origin, transit and destination.

5. Post-intervention measurement and analysis

After the intervention is made, a number of steps must be taken to successfully produce an effective impact assessment.

- First, there must be some sort of post-intervention measurement. The potential types and methods of this are the same as outlined in Step 3.

- Second, the impact must be calculated: ideally, the level of the variable of interest as measured after the intervention subtracted from the level of the variable of interest prior to the intervention. However, as already discussed, where that is not possible, some other form of baseline can be used, as drawn from secondary data or a constructed counterfactual. Regardless of the method, a 'bottom line' impact should be calculated, with theoretical assumptions and caveats regarding methodological shortcomings and trade-offs (which are, in all cases, unavoidable) openly acknowledged.

- Third, even when using the most robust research methods, some theoretical discussion of why the observed impact took place is a necessary component of any impact assessment, and usually the most interesting one: Was it purely the intervention or other factors that might have changed pre- and post-intervention? How can we tell? Which elements of the intervention are likely to have had which effects? How can we isolate these? Is this intervention universally applicable? Would it work on a different target audience? Why might the effect be different? What about the context might have affected the bottom line of the intervention? Naturally, as the robustness of the methodology decreases (e.g. comparative case studies), the need for greater discussion of these issues increases.

- Fourth, with the causal mechanism identified, recommendations for future communication campaigns should be created or, if this is a continuous activity, campaign or project, changes can be made based on these findings.

With regard to migration communication campaigns specifically, Trajden (2018: 7) find that 'the majority of the campaign evaluations claimed that the campaign under study was "successful" in inducing a change

in knowledge, perceptions, attitudes and – to a lesser degree – (intended) behaviour. However, most of the evaluations reviewed provided relatively little evidence of the full impact of information campaigns. While many of the evaluations reported the number and profiles of campaign recipients or beneficiaries, impact was not directly measured.' As such, it seems safe to conclude that most migration IAs do not calculate a final bottom line of the impact.

Moreover, Macnamara (2020b: 436) gives a thorough explanation of why the final calculation of the bottom line impact should be in terms of outcomes, which should derive fairly directly from the stated objectives of the intervention and why some theoretical consideration of why the observed impact took place: 'Numerous studies have shown that evaluation in PR, corporate communication, strategic communication, and related fields of practice is predominantly focused on outputs, rather than outcomes or impact of communication [...] contemporary approaches shift the focus of evaluation to outcomes and impact.' This is the result of the pervasive idea that outputs directly and predictably result in outcomes: 'Preoccupation with outputs with comparatively little critical attention paid to outcomes and impact is largely based on lingering assumptions about media effects [... however ...] As many scholars point out, direct effect theories of media and communication have been dismantled [...] in favor 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. Crucially, whereas measurement only requires data, evaluation (which includes impact assessment) requires some theoretical consideration of why the impact took place. Despite professionals in advertising, PR, government communication, strategic communication, and communication management both focussing more on measurement and confusing measurement and evaluation, the greater value lays in understanding why whatever change there is happened—this allows us to form recommendations for future interventions ... While measurement (the taking of measures) is necessary and a precursor to evidence- based evaluation, the latter involves "making a judgment" about the value or significance of something.' As such, IA is incapable of producing long-lasting lessons, recommendations and, ultimately, progress. without theoretical consideration of why there was an impact.

Indeed, this is particularly relevant to migration communication. As Trajden (2018: 13) states: 'the lack of such evidence leads many observers to call the general approach taken by information campaigns into question [...] For implementers of information campaigns, there is a crucial need for guidance on how to design such campaigns in order to achieve certain desired effects on the awareness, knowledge, perceptions, attitudes and (intended) behaviour of the specific target groups.' Again, this requires theoretical arguments about why certain interventions had the impact that they did. However, regarding campaigns dissuading irregular migration, 'Generating evidence in this area is difficult, given the complexity and

variety of migration information campaigns and their underlying assumptions [...] First, it is assumed that potential migrants lack information; second, that available information (i.e. prior to an information campaign) is inaccurate; third, that new information (i.e. from the campaign) is trusted and believed; fourth, that the new information will affect knowledge, perceptions and/or attitudes; and, fifth, that a change in knowledge, perceptions and/or attitudes will translate into a change in behaviour.' This provides an excellent example of, first, why being open about one's theoretical assumptions is necessary and, two, why challenging those assumptions is the route to progress.

More practically, the UK GCS (2018) recommends a number of post-intervention steps beyond simple measurement of impact. The most relevant of those of UK GCS (2018: 6) state:

- 'It is recommended that approximately 5 to 10% of total campaign expenditure is allocated to evaluation. In addition to operational data, evaluation costs will often include commissioning research to measure awareness and message penetration levels.'

- 'A comparison of actual outcome data with targets set in objectives. Were the objectives met? If not, what reasons can be offered to explain the variation? If the objectives were surpassed, what has driven that?'

- 'Considering the causal link between the subject-oriented outtakes and the outcomes. Some campaigns will be more effective in converting awareness and attitudinal changes into tangible behavioural outcomes.'

- 'Findings for current or future campaign optimisation. Ideally this will include attribution modelling and econometric analysis (scientifically assigning a proportion of 'cause' to different elements, messages and channels of a campaign). Even without advanced studies, campaigners can often draw conclusions about which channels have been more or less effective than anticipated. Is there anything that others can learn from your theory of change?'

- 'A conclusion including whether or not the campaign was successful in achieving its policy aims. This should also include what we would do differently next time or for future similar campaigns.'

Finally, Gertler et al (2016: 320) offer the following similar points, amongst others, in checklist form that all IA should include, again emphasising the importance of theoretical considerations:

- 'A concrete and relevant policy question—grounded in a theory of change—that can be answered with an impact evaluation.'

- 'A robust methodology, derived from the operational rules of the program, to estimate a counterfactual that shows the causal relationship between the program and outcomes of interest'

- A methodology and sample that provide results generalizable for the population of interest'

- 'Impact evaluation results should be informed by complementary process evaluation and

monitoring data that give a clear picture of program implementation. When programs succeed, it is important to understand why. When programs fail, it is important to distinguish between a poorly implemented program and a flawed program design.'

Conclusion

The demand for assessing the impact of communication campaigns in the field of migration is increasing. However, migration professionals have reported significant difficulties in performing such assessments—not least because of the, often unrealistic, methodological standards of academia or international organisations—and had their attempts at such assessments severely criticised. This report brings together disparate terminology, findings and recommendations from the private and public sectors and academia to synthesise a set of five general steps for practitioners when performing IAs in the migration sector. The overriding argument across the five steps is that, rather than being a necessarily complex and expensive exercise as suggested by current guides, IA simply involves comparing a variable of interest *after an intervention to what it would be if there had been no intervention.* Given the hypothetic nature of the comparison, the results are bound to be qualified and contingent and, once migration practitioners understand the theoretical and methodological considerations resulting from this, the quality and quantity of impact assessments in migration policy should increase, potentially a source of significant progress in the sector.

References

Adelle, C. and S. Weiland. 2012. Policy Assessment: the State of the Art. Impact Assessment and Project Appraisal, 30:1, 25-33, DOI: 10.1080/14615517.2012.663256

AMEC. 2020. Barcelona Principles 3.0. Accessed at https://amecorg.com/resources/barcelona-principles-3-0/ on 4 August 2020

Bamberg, K. 2019. 'Moving beyond the 'crisis': Recommendations for the European Commission's communication on migration'. European Migration and Diversity Programme. Discussion Paper. Accessed at https://www.fes-europe.eu/fileadmin/public/editorfiles/events/2019/Q4_2019/Moving_beyond_the_crisis.pdf

Banulescu-Bogdan, N. 2018. 'When Facts Don't Matter: How to Communicate More Effectively About Immigration's Costs and Benefits' Migration Policy Institute. Accessed at https://www.migrationpolicy.org/research/when-facts-dont-matter-immigration

Browne, E. 2015. 'Impact of communication campaigns to deter irregular migration' GSDRC Helpdesk Research Report. Accessed at https://assets.publishing.service.gov.uk/media/57a0896840f0b652dd0001f4/HQ1248.pdf on 1 August 2020.

Christiano, An. 2017. 'Foreword: Building the Field of Public Interest Communications'. Journal of Public Interest Communications, 1(1): 4-15

Dennison, J. (2019, online first) 'How issue salience explains the rise of the populist right in Western Europe' in International Journal of Public Opinion Research

Dennison, J. and Geddes, A. (2019) 'A Rising Tide? The Salience of Immigration and the Rise of Anti-Immigration Political Parties in Western Europe' in The Political Quarterly, 90(1): 107-116

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

European Commission, 2017, 'TOOLKIT for the evaluation of the communication activities' DIRECTORATE-GENERAL FOR COMMUNICATION. Accessed at https://ec.europa.eu/info/sites/info/files/communication-evaluation-toolkit_en.pdf on 31 July 2020

European Migration Network (2012). Practical Measures to Reduce Irregular Migration. Synthesis Report. European Migration Network http://extranjeros.empleo.gob.es/es/redeuropeamigracion/Estudios_monograficos/EMN_Synthesis_Report_Irregular_Migration_October_2012.pdf

Gertler, P.J. et al. (2018). Impact Evaluation in Practice. Second edition, Inter-American Development Bank and World Bank, Washington, D.C.

Gregory, A. 2020. 'The Fundamentals of Measurement and Evaluation of Communication' in, eds., Luoma-aho, V. and Canel, M-J, The Handbook of Public Sector Communications. John Wiley and Sons, pp. 367-382

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.

Henningsen, A., Traverse Healy, K., Gregory, A. Johannsen, A., Allison, R., Bozeat, N., & Domaradzki, L. (2014). Measuring the European Commission's communication: Technical and methodological report. http://ec.europa.eu/dgs/communication/about/evaluation/index_en.htm.

Macnamara, J. 2020a. 'Public sector communication measurement and evaluation' in, eds. Luoma-aho, V. and Canel, M-J, The Handbook of Public Sector Communications. John Wiley and Sons, pp. 361-365

Macnamara, J. 2020b. 'New Developments in Best Practice Evaluation: Approaches, Frameworks, Models, and Methods' in, eds. Luoma-aho, V. and Canel, M-J, The Handbook of Public Sector Communications. John Wiley and Sons, pp. 435-454

Marthouz, J-P. 2006. 'How to communicate: Strategic communication on migration and integration'. King Baudouin Foundation. Accessed at http://www.media-diversity.org/en/additional-files/documents/A%20Guides/How%20to%20Communicate_Strategic%20communication%20on%20migration%20and%20integration%20%5bEN%5d.pdf

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

Noblet, A., & LaMontagne, A. (2009). The challenges of developing, implementing, and evaluating interventions. In S. Cartwright & C. Cooper (Eds.), The Oxford handbook of organizational well-being (pp. 466–496). Oxford, United Kingdom: Oxford University Press.

O'Neil, G. (2020) 'Measuring and Evaluating Media: Traditional and Social' in, eds., Luoma-aho, V. and Canel, M-J, The Handbook of Public Sector Communications. John Wiley and Sons, pp. 405-416

Romenti, S. and G. Murtarelli (2020) 'Measuring and Evaluating Media: Traditional and Social' in, eds., Luoma-aho, V. and Canel, M-J, The Handbook of Public Sector Communications. John Wiley and Sons, pp. 383-403

Sharif, H. 2019. 'Communicating effectively on migration: recommendations and policy options'. Resoma: Research Social Platform on Migration and Asylum. Policy Option Brief: December 2019. Accessed here: https://www.migpolgroup.com/wp-content/uploads/2020/01/POB-Public-Opinion-Hind-Sharif.pdf

Thomas, J. (2008). Advertising effectiveness. Decision Analyst [Web site]. Retrieved from http://www. decisionanalyst.com/publ_art/adeffectiveness.dai

Tjaden, J., S. Morgenstern and F. Laczko (2018), "Evaluating the impact of information campaigns in the field of migration: A systematic review of the evidence and practical guidance", Central Mediterranean Route Thematic Report Series. International Organization for Migration, Geneva.

UK Government Communications Service. 2018. Evaluation Framework 2.0. Accessed at https://3x7ip91ron4ju9ehf2un-qrm1-wpengine.netdna-ssl.com/wp-content/uploads/2020/03/Evaluation-Framework-2.0.pdf on 4 August 2020

Welcoming America's (2018) 'Stand Together: Messaging to Support Muslims and Refugees in Challenging Times'. Accessed at https://www.welcomingrefugees.org/sites/default/files/documents/resources/Stand%20Together%20Toolkit.pdf on

WHO, 2017. 'WHO Strategic Communications Framework for effective communication' Accessed at https://www.who.int/mediacentre/communication-framework.pdf?ua=1 on 31 July 2020.

Contact:
EMM4_team@icmpd.org

Address:
ICMPD Regional Coordination Office for the Mediterranean :
Development house, 4A / St Anne Street / Floriana, FRN 9010 / Malta tel:+356 277 92 610

www.icmpd.org/emm4

emm4_team@icmpd.org

@EMM4_migration