

Resources to support the gambling industry to do evaluation

Why evaluate?

Action to minimise gambling-related harm is essential, but action alone is not enough - evaluation tells us what's working and what is not. It enables informed decisions about what interventions to put in place and how these can be improved.

One of the five top priorities of the National Responsible Gambling Strategy is:

“To build a culture where new initiatives are routinely evaluated and findings put into practice”.

The Responsible Gambling Strategy Board, supported by the Gambling Commission and GambleAware, has put in place an evaluation protocol. This protocol provides the framework for industry to use when evaluating their interventions. It is based on four principles: robustness and credibility; proportionality; independence and transparency.

What are these resources?

The resources provided here have been put together to demystify evaluation, and what it can do for you, and provide practical guidance on using evaluation, and using it well.

Some of the more common questions you may ask yourself are set out as FAQs. These are organised against the four principles in the evaluation protocol and provide more information about these principles. These are backed up by some practical resources and tips which can take you further. If you want to go still further, we have also provided some links to additional sources. There are case studies to provide examples of the use of evaluation in practice.

Who are the resources for?

This guidance is for anyone setting up or trying to improve a responsible gambling or harm minimisation initiative, action or interventions. This could be an organisation in or supporting the industry, sector and trade bodies, operators, treatment and education providers, etc.

They are 'starter' resources. If you are going to do your own evaluation they provide pointers on planning, options and opportunities; if you are to commission an evaluation they will help you become an informed client better placed to specify, commission and steer what is to be done. They do not give specific 'how to' guidance on detailed methods but provide enough to get you started on asking yourself, and others, the right questions – and to answer them.

The resources are aimed at:

- Executives in operators who want to get to grips with why and how evaluation is needed in shaping harm minimisation
- Managers who are running trials or pilots and who need practical ways of providing trusted evaluation evidence to inform decisions on roll-out
- Corporate Social Responsibility teams, and others, looking to understand how to interpret and use evaluation evidence in shaping policy and practice
- Education and treatment services who need to use evaluation to improve the effectiveness of their support, and demonstrate what works and how.

Where next?

We realise that different people, in different parts of the sector, will have different needs. Gamble Aware is keen to have feedback on the usefulness of these resources, and any suggestions for further support, or next steps, which can be sent to natalie@gambleaware.org.

FAQs

Question	Answer
	P1: Robustness and credibility
<i>Why do harm-minimisation initiatives need 'robust' evaluation?</i>	There are two very good reasons. Firstly, it makes good business sense – how else do you demonstrate convincingly to others that something is working, worthwhile and merits the investment put into it. Secondly, the regulator and licensing authorities have a keen interest in evidence that harm-minimisation interventions work.
<i>What can evaluation do that common sense cannot tell us?</i>	Common sense and opinion has a value but it is a poor substitute for numbers, hard evidence, customer feedback and insights and independent analysis. This is especially when you need to show that something is delivering what you expected and is working well (or is not). Opinion without this sort of evidence will always risk lacking in credibility and will be open to challenge, especially by sceptics and those holding purse-strings.
<i>I am surrounded by data about what we are doing and what it costs; what can evaluation add to that?</i>	<p>Available data is a good start. In areas like machine-based play it may even provide for all (or most of) what you need. But data still needs collating and making sense of, and independent evaluators will provide for a more credible and trusted assessment. Available data is often not sufficient – especially on understanding outcomes and impact, not just outputs and process.</p> <p>Harm-minimisation is aimed at making a difference to players (and staff). If you are only able to talk about what outputs or participation an intervention has, it will fall short of showing if it works – if it makes a difference.</p>
<i>When is the right time to start preparation for an evaluation?</i>	<p>As soon as possible and its always best if an evaluation is planned alongside the intervention itself.</p> <p>You may even need to start evaluation before an intervention gets underway – what are called <i>ex ante</i> evaluations can be used to help plan an intervention and/or to provide for 'baseline' information which will be used later to see how much difference the intervention has made. The case study a gambling prevalence <i>ex ante</i> evaluation shows how useful these can be.</p>

<p><i>What types of evaluation should I be using?</i></p>	<p>Evaluation is done to help make decisions about things like cost-effectiveness, impacts, transferability (can a pilot be rolled out?). So, the type of evaluation you need will depend on what you are evaluating and how you need to use the evidence. This comes down to four choices – process, economic, impact and plural evaluations.</p> <ul style="list-style-type: none"> • A process evaluation – which evaluates the mechanisms through which an intervention takes place, its outputs (not outcomes) and effectiveness. • An economic evaluation – which evaluates the costs of inputs, outputs or outcomes or overall value of an action. • An impact evaluation – which evaluates intervention outcomes or longer term impacts (the consequential changes resulting from an intervention). • A plural evaluation – evaluations which combine two or more of these approaches. <p>Resource A sets out some of the things you will need to think about in making the right choice.</p>
<p><i>Is a cost-benefit review the same as an economic evaluation; will I need special expertise to make it credible and robust?</i></p>	<p>Economic evaluations are based on principles of cost-benefit analysis. So, to do them needs an understanding of applied economics, but not all economic evaluations have to be complicated. Some may be quite straightforward and looking only at costs ('cost-description' evaluations) or cost-effectiveness, where the complexity will depend on the nature of what's being evaluated, its expected inputs and outputs and possibly also outcomes.</p> <p>Cost-benefit evaluations are the least straightforward and even for relatively simple harm-minimisation initiatives are likely to be highly complex and will almost certainly need specialised evaluators to provide robust evidence.</p>
<p><i>How do I handle the ethical side of an evaluation?</i></p>	<p>Ethics in evaluation is complex but for industry harm-minimisation evaluations it mainly concerns the way new evidence is collected and used. Crucially, it needs to ensure that what and how an evaluation is done does not place anyone involved at undue risk of harm. Research especially with vulnerable people such as problem or at-risk gamblers, need to be able to show that their arrangements for selecting participants, briefing them (and securing informed consent to take part), collecting evidence, storing and reporting it, are ethically sound.</p> <p>All evaluations should be subject to an ethical review and many will need to go through a formal ethical clearance process, especially if it involves some external funding. External evaluators will be able to advise you on ethical compliance and clearance. A good starting point is the UK Social Research Associations guidance (see pp.25-40 in particular): http://the-sra.org.uk/wp-content/uploads/ethics03.pdf</p>
<p><i>Is it possible to do a robust impact evaluation when it is not straightforward to define impacts?</i></p>	<p>Yes. The challenge is in first defining a small number of appropriate and measurable impacts. A good impact evaluation will have a sharp focus on what is relevant and possible to measure and understand. It will often combine both <i>hard impacts</i> (e.g., lower levels of player debt) and <i>soft impacts</i> (e.g., player awareness of risk behaviours). They may also look out for 'indirect' impacts – unexpected consequences or effects from the initiative.</p>

<p><i>OK but aren't some impacts impractical to get to grips with, such as minimised harm?</i></p>	<p>No. Harm-minimisation and responsible gambling are certainly challenging to reduce to a handful of things to be measured, but it can be done by looking at what the initiative is about, what it is delivering and what is expected to change as a result (in both short term 'outcomes' and longer term 'impacts').</p> <p>It is important to start off with a model of such expectations – sometime called a 'theory of change' which is a great tool not only in planning the intervention and its focus but also an aide to defining what aspects of harm-minimisation or behaviour change need to be assessed by evaluation.</p>
<p><i>What is counterfactual analysis and how do I do it?</i></p>	<p>It is one thing to measure change occurring, for example, to players or participants in an RG intervention; quite another to assess how much of that change resulted directly from taking part (and not other influences). The 'counterfactual' uses tried and tested methods to measure what would otherwise have happened if the initiative had not taken place at all. Resource B sets out what some of these methods are, and where they are best fitted.</p> <p>Evaluation results will not be robust or credible if they cannot say what contribution the initiative is likely to have made to the measured changes. Randomised control trials – RCTs – are often said to be the 'Gold Standard' for impact evaluation, but other counterfactual approaches include 'quasi-experimental' and 'non-experimental methods. RCTs will rarely be relevant to the majority of RG evaluation but quasi-experimental and non-experimental can often provide all you need to show what's working and how well. Resource C provides a ready reckoner tool for the most robust methods – RCTs and quasi-experimental.</p>
<p><i>What role does qualitative evidence play?</i></p>	<p>In general, qualitative methods such as case studies, beneficiary or customer interviews or focus groups, are a valuable part of process evaluations. They are also often a key part of impact evaluations where they can say things about how and why interventions work which relying on quantitative methods alone may not. Only in RCTs are they are difficult to combine with quantitative evidence. Resource A also gives some ideas of where they best fit in.</p>
	<p>P2: Proportionality</p>
<p><i>How do I manage for realistic expectations of evaluation?</i></p>	<p>A starting point for any evaluation is having a very good handle on what can be realistically be expected of it, set out in aims and objectives. This comes ahead of any choices on methods and design. It needs those setting the terms of the evaluation to be both clear about what is needed for the evaluation and to agree what is realistic for it to do. Resource D sets out a handy tool – the R-O-T-U-R framework – for doing this. And remember, any expectation which is not clear or is unrealistic needs to be renegotiated or recognised as going further than the evaluation can accommodate.</p>
<p><i>How do I make an evaluation proportionate?</i></p>	<p>Most harm-minimisation evaluations do not need to be overly technical and the task is about deciding what is possible and practical. This is about making an evaluation 'proportionate' to needs and circumstances. This involves taking account of the state of play of an intervention (whether it is a trial, pilot or running for some time), its level of innovation, its complexity (maybe it involves several inter-related activities not just one), and how long it is to run for before the</p>

	evaluation concludes. Resource E provides a checklist for some of the factors to balance in making your evaluation proportionate.
<i>What should a good evaluation cost?</i>	<p>There is no straightforward answer to this often asked question; and no ready yardsticks. A good evaluation will be proportionate, led by needs and cost effective but its actual budget will depend on many things, including if it is being done internally or externally.</p> <p>The evaluation budget may already be fixed before a design is put in place, so the evaluator will either need to manage down expectations to work inside the budget, or convince the budget holder to spend more. It is best for the evaluation to be anticipated when an intervention is planned and with a separate cost heading for it allowed in the overall budget.</p> <p>When specifying an external evaluation, it may be best to indicate a broad cost range (under £25,000; £60-90,000; etc.) rather than a specific budget and encourage bidders to put in any added-value options if they can justify these. This is a good way to get best value (but not usually at lowest) cost.</p>
	P3: Independence
<i>What is needed for an evaluation to be independent?</i>	An independent evaluation is able to demonstrate management of conflict of interests and impartiality in how it interprets evidence and reaches conclusions. In practical terms, independent evaluators may well be familiar with agencies/companies or interventions that they are evaluating, as long as they have not been involved in its pre-evaluation planning or implementation or otherwise have a stake in its outcomes. An independent evaluation will also need to ensure that the working arrangements with the client and the practical management and steering of the evaluation upholds the evaluator's impartiality and independence of judgment.
<i>When is it best to conduct an evaluation in-house or to commission outside evaluators?</i>	<p>An internal evaluation will always face challenges in demonstrating independence because those conducting will be seen as having an interest in the success (or failure) of the initiative. Steps can be put in place to separate the internal evaluators from the intervention delivery but their judgements will still risk being seen as being 'compromised' by being part of the delivery agency or company.</p> <p>If what is being evaluated is sensitive or controversial, or where findings will be met by stakeholders with pre-set opinions or by doubters; it always best to conduct an evaluation externally, through a procurement and management process which can demonstrate its impartiality. Resource F sets out some of the pros and cons of internal and external evaluations.</p>
<i>Where do I go for a competent evaluator</i>	Choosing a reliable evaluator is usually the most important decision for any independent evaluation. Even if you have a structured procurement process to follow you can give this a helping hand by making sure some expert and established specialist evaluators know about the tendering process and its timing. You will need someone who can show they are

<p><i>who can do the job on time?</i></p>	<p>not conflicted, with a track record of systematic evaluation, an appropriate mix of quantitative and qualitative methods, and who can prove past delivery and credible and comprehensible reporting.</p> <p>Be wary of picking gambling sector specialists or consultants who may have research skills but are not genuine experts in evaluation. Gamble Aware has its own list of evaluation specialists who might provide a starting point.</p>
	<p>P4: Transparency</p>
<p><i>What is a transparent evaluation?</i></p>	<p>Transparency is important to evaluation because, like independence and impartiality, it helps to build confidence and credibility in findings and conclusions. This means evaluations should be as open as possible, through the whole process – the intervention rationale, evaluation objectives and who is funding it; who is doing the evaluation and how they were selected; the evaluation plan, methods and data; as well as sharing the results and conclusions (and their limitations).</p> <p>Transparent evaluation also analyses and sets out in reporting both its <i>reliability</i> (the quality and strength of the methods used) and <i>validity</i> (of the evidence and the generalisability of the findings).</p> <p>Commercial considerations might limit some aspects of transparency, but the more open and transparent an evaluation can be the more it is likely to increase confidence and credibility. For some evaluations transparency may also mean respect for third-party interests and clarity about any rights worthy of protection.</p>
<p><i>How does an evaluation 'constructively' engage to help build confidence and credibility?</i></p>	<p>For an evaluation to constructively engage it needs to be as open as possible, and accessible to those who might want to ask questions of why it's being done and how, at all stages. It is easier for an evaluation to get on with the job and leave external dialogue to once the final report is wrapped up, but to do so may miss opportunities for evidence sharing and critical review and may increase suspicion about the evaluation or evaluators.</p> <p>Just how much constructive engagement is possible depends on the nature of what is being evaluated, stakeholder relationships and expectations and issues of data or commercial sensitivity. But within most evaluations there are many opportunities to engage outside the evidence collection. This does not mean changing direction or methods because a stakeholder ask for it; but it may mean explaining why some things cannot be done and what's being done instead. Well managed this takes nothing away from evaluation independence and impartiality and adds a lot to credibility and confidence. Resource G sets out some of the many possibilities.</p>

Jargon Buster:

A short glossary of selected evaluation terms

Evaluation is full of technical terms. Many are for quite common sense ideas. Those set out here should provide a useful starter guide to unpicking what is meant.

Additionality: This is the change or impact measured or observed from an evaluation of an intervention which is over and above what was expected.

Attribution: A finding from an impact evaluation which shows just how much the intervention itself was itself responsible for the outcomes and impacts being measured [see also causality and counterfactual below].

Before and after analysis: A simple (non-experimental) method which helps to estimate attribution by contrasting outcomes during an intervention (and perhaps at the end of a pilot or trial) with data before the intervention took place.

Blinded evaluation: A blind, or blinded, evaluation is where information about the test is masked from the participant and others until after the evaluation outcome is known. This is an important part of a Randomised Control Trial and ensures the results cannot be biased by (inadvertently) distorting the behaviour of people participating in or otherwise involved the trial.

Causality: A finding or observation from an evaluation of an intervention which digs deeper than looking at the 'overall' (gross) impacts and measures or estimates that part of the gross impact which can be directly attributed to the intervention itself – the 'net' impact. In other words, a 'causal' analysis separates out (discounts) the contributions of the intervention itself from any other (e.g., external) contributions to the impacts achieved. *NB. One reliable way of doing this is for the evaluation to develop what is called a 'counterfactual' analysis or case – as below.*

Comparative group: A method in quasi-experimental (impact) evaluation often used instead of a Randomised Control Trial, and which contrasts the measured outcomes in an intervention area (or group of people) with a very closely matched *comparison* group (e.g., a like-for-like geographical area). The contrast can be used to demonstrate causality or the added-value or additionality of the intervention.

Control group (and analysis): A method of impact evaluation used in Randomised Control Trials which assesses causality of impacts by contrasting the results for the beneficiaries or participants (of an intervention) with a closely matched, randomly selected, 'non-intervention' or control group.

Counterfactual analysis: An analysis as part of an impact evaluation which sets out to identify what would have occurred if an intervention or activity had not been implemented and comparing this to the measured outcomes after the intervention.

Control groups (in an Randomised Control Trial) or comparison groups (in a quasi-experimental evaluation) are reliable ways of doing this.

Deadweight: An identified impact or benefit (or part of it) from an intervention which the evaluation shows would have happened even if the intervention had not taken place – a ‘deadweight’ effect.

Gross impact: An overall (non-attributed) outcome or impact resulting from an (evaluated) intervention or activity (see impact below).

Hybrid evaluation: An evaluation methodology using mixed methods – and typically combining quantitative and qualitative methods to contrast and triangulate (see below) different evidence sources.

Impact: An observed effect resulting from an (evaluated) intervention and as a consequence of delivering or achieving specific activities or ‘outputs’. This is usually associated with measuring medium or longer-term changes (e.g., sustained behaviour changes) and which may take some time to be realised; outcomes (see below) refer to shorter term changes.

Knock-on impact: An unexpected, unintended or indirect consequential effect of an (evaluated) intervention (see impact above).

Leakage: Effects within measured impacts which support others outside the targeted or expected intervention group.

Monetised: The process which results in an outcome or impact being converted or translated into a quantified cash or financial value.

Net impact: An outcome or impact attributed to a specific intervention or activity which discounts changes which would have otherwise have occurred without the (evaluated) intervention or activity having taken place.

Opportunity cost: A benefit, profit, or value of something that must be given up to acquire or achieve something else. *NB Economist use this to assess the real return on an investment or intervention and refer to it as the next best alternative foregone.*

Outcomes: A short term effect resulting from an (evaluated) intervention and usually resulting as an early consequence of delivering or achieving specific activities or ‘outputs’ (see also impact above).

Primary evidence: Quantitative and/or qualitative evidence in an evaluation which is generated directly by the evaluator (or on their behalf) from additional information collection methods.

Proportionality: The principle of evaluation design which sets out that in addition to the need for reliable information, the choice and mixture of evidence gathering and analytical methods to be used should be ‘proportionate’ to the objectives, scale and nature of the programme being evaluated.

Secondary evidence: Quantitative and/or qualitative evidence in an evaluation which is collated from existing sources of evidence within or outside an intervention including from, for example, management or monitoring information and diverse documentary sources.

Substitution: Measured outcomes or impacts (or aspects of them) on an intervention group which are realised at the expense of others outside the intervention group, often as unintended consequences from the intervention (see below).

Triangulated evidence: Triangulation is a commonly used approach in all forms of evaluation that provides for validation of both quantitative and qualitative evidence through cross verification from two or more sources, typically derived from combination of several research methods in assessing the same phenomenon.

Unintended consequences: Unexpected impacts and effects of (evaluated) interventions and activities which need to be identified and taken into account in any assessment of net impacts.

Valuation: Techniques for measuring or estimating the monetary and/or non-monetary value (see above) of observed outcomes and impacts, contributing to an assessment of added value or cost-effectiveness of the evaluated intervention.

Value for Money: Value for money (VfM) measures the extent to which an intervention (or sets of activities) has provided the maximum benefit for funding bodies from the resourcing of activities, benefits secured and outcomes and impacts arising. VfM provides a quantitative measure, typically for specific goods or services, or combinations of these.

Process and Impact Evaluation of Gambling Prevalence and Harm-Minimisation Support in Leeds City Region

The 'Gambling Prevalence' evaluation, was commissioned by Leeds City Council in advance of the opening of the GGV 'super-casino' in Leeds City Centre. The ex-ante evaluation provided a baseline of gambling prevalence across the wider Leeds metropolitan area and an initial review of referral and support services for problem gamblers and those at risk.

What was the evaluation about? Leeds City Council (LCC) licenced Global Gaming Ventures (GGV) to open a large casino as part of the Victoria Gate city centre redevelopment; this was to be the 4th largest casino in Britain. The evaluation was commissioned following a public consultation raising some concerns about an increased risk of problem gambling, and insufficient support services. The evaluation was to provide independent evidence to guide a city-wide and multi-stakeholder harm-minimisation strategy.

When was it done, who by? The four-month evaluation started in April 2016. It was commissioned from a team led by Prof David Parsons and Dr Alex Kenyon of Leeds Beckett University, following a national competition open to all bidders. Team members involved Dr Heather Wardle to provide continuity with other harm-minimisation research nationally.

How was it done? The intensive evaluation needed to make best use of available national and local data, supplemented by evidence from operators, support providers, referral and other agencies, as well as 'at risk' gamblers. A large scale resident-based 'prevalence' survey would have had data collection and validity challenges, budget and time constraints. Instead the team put together a prevalence analysis based on DSMV IV and PGSI data from the National Gambling Prevalence Survey (NGPS) and the Health Survey for England (HSE) using sub-regional samples, regional and national data. The evaluation combined:

- A 'Quick Scoping' rapid evidence review to update the review *Exploring Gambling-related harm: who is vulnerable?* National study (2015) and assess Leeds implications.
- Prevalence analysis based on Leeds and 8 comparator areas (selected by matching socio-economic and demographic profiles from official DEFRA and other data sets).
- Additional prevalence and trend data was collected from sector bodies and sampled retail operators.
- Further evidence from a stakeholder consultation and semi-structured interviews with 20 stakeholder organisations, including generic, specialist and gambling specific support bodies in the city-region. Interviews were also conducted with carefully selected 'at risk' gamblers after an ethically-controlled, operator and stakeholder-based recruitment process.

Interim analyses were prepared and considered at three meetings of a cross-stakeholder steering group chaired by LCC, as were a draft and final report and recommendations. The report was launched in March 2017 following LCC further stakeholder discussions on ways forward.

What worked well? The evaluation team worked independently in evidence collection, analysis and recommendations, but in close collaboration with LCC. The evaluation showed how much could be achieved in setting baselines for any future evaluation from sophisticated comparative analysis using existing data sets, combined with a wider 'qualitative' engagement of stakeholders. Reporting to a large multi-stakeholder steering group raised local confidence in the analysis. This also helped with access to local agencies

(only two refused to take part), and added to the credibility of the LCC action plan developed following the evaluation.

What lessons were learnt? The evaluation provided a comparative review of gambling prevalence in Leeds which can be harnessed for monitoring any changing features, trends and support needs. It identified gaps in referral arrangements and between demand and provision for specialist services, and provided recommendations for addressing those gaps. It also showed that the inter-relationships between retail operators (and between retail and online providers) means that any future impact evaluation of changes in gambling-related harm needs to be conducted across gambling sectors and not be limited to a single operator, such as the new casino.

Further information: The evaluation report is available at:
<http://www.leeds.gov.uk/docs/Problem%20Gambling%20Report.pdf>.

Process Evaluation of the Player Awareness System pilot

The Player Awareness System (PAS) pilot involving six retail operators was evaluated in 2016 through an independent analysis commissioned by the (then) Responsible Gambling Trust (RGT). This was a process evaluation, conducted by PriceWaterhouseCoopers (PwC) through large data-set analysis across the participating operators. The findings led to substantial improvements in PAS algorithms, implementation and monitoring, and a commitment to a further evaluation to assess improved effectiveness in harm-minimisation.

What was the evaluation about?

Research for RGT in 2014 encouraged the industry to identify markers of harmful gambling in order to intervene with players at risk. Using this and other research, the Association of British Bookmakers (ABB) and members developed Player Awareness Systems (PAS). These use behavioural indicators to distinguish between problem and non-problem gambling, by people using accounts to play gaming machines in licensed betting offices. The aim of PAS is to intervene to prevent customers becoming problem gamblers, by spotting those who are on a trajectory towards harmful play and intervening with messages in order to make the player aware of their own behaviour and to halt and reverse that trend. This innovative initiative directly involved Coral, Ladbrokes, Paddy Power, William Hill and machine manufacturers Scientific Gaming and Inspired Gaming (on behalf of independent bookmakers), with all ABB member signed up to a wider roll-out.

When was it done, who by? To ensure its independence, the management of the evaluation was undertaken by RGT and externally commissioned following an open tendering process. The contract was awarded to PwC who acknowledged its existing auditing relationship with Ladbrokes (one of the pilot operators). This was not regarded by the RGT selection panel as a conflict of interest. The evaluation commenced in March 2016.

How was it done? PAS implementation was at an early stage, with the pilot scheme put in place early in 2016. This involved each operator developing different algorithms and ways of intervening. Process evaluation was needed to review early implementation, compare effectiveness across the different PAS processes, identify improvement needs and provide insight for subsequent development. This was a multi-method evaluation combining an intensive quantitative analysis with operator interviews, specifically:

- System and algorithm review to compare the different systems
- Analysis of machine data with sample testing of whether messaging and customer interactions were delivered as specified and there were adequate internal controls
- Interviews with operators and systems designers to review controls, system responses (e.g., customer/staff alerts), implementation challenges and system gaps.

This early evaluation was limited to customers whose sessions could be tracked through loyalty cards. The evaluation did not review costs or impacts – both of which were seen as premature given the pilot nature of the scheme.

What worked well? The evaluation was conducted across multiple systems with different characteristics. This comparative aspect of the evaluation provided a rich source of evidence of which aspects of operator arrangements worked better or less well. Undertaking a process evaluation early in the implementation helped identify how systems could be improved. The PwC team reported to an RGT steering group which included an external evaluation expert and provided a direct link to operators to ensure access to data and sharing of experience.

What lessons were learnt? The process evaluation showed considerable variety of algorithms used by the ABB's member companies, and differences in maturity. The findings demonstrated PAS was work in progress but showed potential as a systems-based approach to responsible gambling. It also showed scope for closer integration of Fixed Odds Betting Terminal (FOBT) PAS data with data from other operator betting activities, more use of insights from behavioural psychology, and tighter controls for PAS messaging and customer interactions. The process evaluation confirmed it remained premature to look at impacts on player behaviours linked to PAS and suggested a further process review when improvements had been made. This would also assess the most viable focus and timing for a subsequent impact evaluation – to see how PAS encourages customers to think about how they are gambling and influences in changed behaviour.

Further information: RGT/GambleAware has published the full process evaluation at: http://infohub.gambleaware.org/wp-content/uploads/2016/10/PAS-evaluation_Final-report_13102016.pdf

Resource A What are the different types of evaluation (and where do they fit)?

Type of evaluation	Typically for the purposes of ...	Good for ...	Not so good for ...
<p>Process evaluation:</p> <p><i>Evaluating the mechanisms through which a responsible gambling intervention takes place, its outputs (not outcomes), and effectiveness.</i></p>	<ul style="list-style-type: none"> • Providing evidence of how and how well an intervention has been implemented or managed against needs/expectations or budget/targets • Reviewing how it operates, how it produces what it does and differences in effectiveness (e.g., between different operations, types of users, beneficiary social groups or geographical areas) • Identifying improvement potential • Assessing cost-effectiveness and areas for cost-efficiencies 	<ul style="list-style-type: none"> • Accountability (assessing costs against budgets) • Assessing roll-out or scale up potential of a trial/pilot • Understanding 'what works' (and does not), for who/where/when and why • Staged or formative evaluation in a longer-term intervention to inform improvement 	<ul style="list-style-type: none"> • Looking at outcomes or impacts (see impact evaluation)
<p>Economic evaluation:</p> <p><i>Evaluating the costs of inputs, outputs or outcomes or overall value of an action.</i></p>	<ul style="list-style-type: none"> • Measuring costs and cost-efficiencies against business plans/budgets or other expectations • Quantifying cost-efficiencies and cost-effectiveness in money terms • Measuring or estimating value for money or value-added of initiatives 	<ul style="list-style-type: none"> • Accountability (assessing costs against budgets) • Projecting cost-efficiencies or cost-utility • Reviewing cost-benefits of outputs or outcomes in money-terms 	<ul style="list-style-type: none"> • Pilot, staged or formative evaluations • Where outcomes cannot be credibly converted to 'money' values

<p>Impact evaluation:</p> <p><i>Evaluating intervention outcomes or impacts (longer term) – the consequential changes resulting from an intervention set against its aspirations.</i></p>	<ul style="list-style-type: none"> • Quantifying outcomes (short or medium-term) or impacts (longer term) resulting from an initiative • Identifying unexpected (additional) impacts or unintended consequences • Assessing the contribution made by an initiative to overall outcomes/impacts (i.e., attribution) • Understanding how different outcomes are occurring, and what influences the impact, including enablers and constraints 	<ul style="list-style-type: none"> • Measuring outcomes and how these come about, for whom and in what circumstances • Setting inputs/outputs against outcomes to assess achievements against expectations • Demonstrating impact to stakeholders 	<ul style="list-style-type: none"> • Interventions lacking clear expectations of impact(s) • Short term evaluations • Interventions without scope/potential for robust quantification
<p>Pluralistic evaluation:</p> <p><i>Evaluations which combine two or more process-economic-impact evaluation approaches.</i></p>	<ul style="list-style-type: none"> • Meeting multiple needs for evaluating an action or intervention, usually over a longer timeframe 	<ul style="list-style-type: none"> • ‘Longitudinal’ evaluation • Long term, sustained intervention and resource commitment • Multi-action interventions taking place in complex environments 	<ul style="list-style-type: none"> • In-house evaluation • Pilot or small-scale interventions • Limited resource evaluations or with short timeframes

Resource B Practical choices in ‘counterfactual’ impact evidence

Evaluation approach	What question are evaluators seeking to answer?	How is the counterfactual assessed?	Possible methods to use	Confidence and likely robustness (and relevance)
Randomised control trial (RCT)	<i>To what extent does the intervention cause the observed outcome(s)?</i>	<p>End of evaluation measured outcome(s) are contrasted for an:</p> <ul style="list-style-type: none"> • Intervention (treatment) group • Pre-determined and parallel ‘control’ group who do not receive the intervention. <p>The sample is randomly selected from a common population, with each member selected by chance and with an equal chance of being selected. The ‘trial’ is carefully controlled to avoid any delivery or external distortions which might affect outcomes.</p>	<p>Fully-experimental with randomised selection through either:</p> <ul style="list-style-type: none"> • Individually sampled participants (an I-RCT), or • Cluster group selections (C-RCT). <p>RCTs use quantitative methods and statistical analyses and can be especially useful for pilots or trail interventions. Qualitative inputs can be added but only where very carefully designed to avoid any risk of bias to the trial.</p>	<p>Very high</p> <p>High</p> <p><i>RCTs are well regarded for quality of evidence but hard to do well, costly and not usually suited to complex interventions such as for harm minimisation.</i></p>

<p>Quasi-experimental (QE) design</p>	<p><i>To what extent does the intervention have the expected outcome(s)?</i></p>	<p>End of evaluation contrast of the intervention outcome(s) for a:</p> <ul style="list-style-type: none"> • Defined group(s) of participants in the intervention, and • Comparative group which is concurrent, closely matched but not randomly selected. <p>Data from the comparative group may be drawn from available sources (where up to date) or by additional research.</p>	<p>Partially experimental - the comparative group might typically come from:</p> <ul style="list-style-type: none"> • Matched (geographical) area • Pre-participation group or area <p>OR</p> <ul style="list-style-type: none"> • Opt-out groups (opt-in are intervention group) • Interventions with groups of people out of scope but which are 'near fit' to use as comparisons with beneficiaries • Interventions with intermittent (on and off) application. <p>QE methods also use quantitative methods and statistical analyses. Unlike RCTs, they can more easily be combined with qualitative inputs such as case studies to better understand how impacts come about.</p>	<p>Moderately high</p> <p>Moderate</p> <p><i>QE can be next best to an RCT for credible evidence but need careful design to provide for suitable comparisons.</i></p>
--	--	---	---	---

<p>Non-experimental</p>	<p><i>To what extent does the intervention make a difference?</i></p>	<p>Outcomes data to contrast against the intervention group is drawn from 'before' and 'after' data for participants or from outside the initiative altogether typically by using external statistical data sets or sources to provide a benchmark.</p>	<p>Non-experimental designs where the comparator evidence may come from:</p> <ul style="list-style-type: none"> • External statistical benchmarks. Benchmark sources may include national survey sources, which can be manipulated to provide as close a comparison as possible to intervention participants • Before (at start of action) and after (at end) contrasts of participants • Participant 'trajectory' analysis using (pre-start) historic data. <p>These are called 'constrained' designs – but non-experimental methods are well suited to combining both quantitative and qualitative methods, to estimate the level and nature of impacts and assess how these come about and why.</p>	<p>Moderate to low</p> <p><i>Non-experimental methods are well suited to interventions where there is a lot of available data to use, and if an RCT or QE design is not possible. Combining more than one of the non-experimental methods helps boost robustness and credibility.</i></p>
--------------------------------	---	---	---	---

Resource C When can you use ‘experimental’ evaluation? – A ready reckoner guide

You may not always be able to use an experimental design for an evaluation – often they will not be possible and other (less robust) methods will need to be used. Budget, nature of the intervention, its context and scale, data availability and likelihood of comparative evidence are the main influences on what is, and is not, possible.

	Randomised control trials (RCT)/quasi-experimental (QE) more viable when:	RCT/QE less viable when:
Budget availability (for evaluation)	<ul style="list-style-type: none"> Internal evaluation: Allocated staff resources/release of suitably skilled and experienced project management and evaluation team External evaluation: Substantial allocated budget to procure appropriate contractors 	<ul style="list-style-type: none"> Limited (or no) allocated staff resource or budget Constraints affecting financing or contracting
Nature of action or intervention	<ul style="list-style-type: none"> New intervention where the need for evaluation is anticipated at design stage Distinct change in system(s) or practice affecting participants ‘Isolatable’ intervention Non-complex and stable intervention environment 	<ul style="list-style-type: none"> New intervention where the need for an evaluation was not anticipated at design stage, or added much later Initiative is already mature or established and being modified Multiple outcome/impact effects expected from initiative High levels of likely unintended consequences (e.g., leakage) Intervention environment not likely to be stable
Expected (likely) scale of initiative impact	<ul style="list-style-type: none"> Large expected effect Distinctive effect relative to other changes taking place 	<ul style="list-style-type: none"> Small expected effect Complex environment (multiple confounding factors)

	<ul style="list-style-type: none"> • Sustainable effect measurable in (relatively) short timeframe 	<ul style="list-style-type: none"> • Long effect lead times or incremental impacts
Anticipated availability of appropriate data and comparator evidence	<ul style="list-style-type: none"> • Appropriate data available on all individual participants • No data access constraints (e.g., data protection) • Data well fitted to intervention period/classification needs • Comparative and control data – before-during-after intervention • Objective selections for intervention/control group • Minimised/measurable selection bias 	<ul style="list-style-type: none"> • Data coverage not comprehensive • Access constraints affecting some/all • Data not adequately differentiated • Data not well fitted to intervention period; data lags • Summative data collection only • Required data limited to pilot area; or inadequate comparisons • Ethical constraints to comparisons • Lack of control group within intervention • Unstructured participant selections

Resource D The ‘ROTUR’ framework for managing evaluation expectations

Evaluation is not just about picking the ‘best’ methods. It’s a real risk of setting off on shaky foundations, or even the wrong path altogether, if it does not set realistic expectations of what is needed, why, who for and when. Managing expectations is about setting solid foundations for an evaluation by paying close attention to: roles and responsibilities; the evaluation’s own outcomes (and deliverables); timing, use and users; and appropriate resourcing. This R-O-T-U-R framework sets out some of the most common ‘do’s’ and ‘do not’s’ for this crucial step of planning.

1. Roles and responsibilities	
DO ...	DO NOT ...
<ul style="list-style-type: none"> Start at the end; who is the end-user (any intermediaries); how/when are they to be engaged in decision-making 	<ul style="list-style-type: none"> Forget to identify internal/external procurement needs (may affect sign off; funding limits; close-open tender; marketing; etc.)
<ul style="list-style-type: none"> Establish who has delegated responsibility for the specification (including objective setting; timetable; resourcing and budget) 	<ul style="list-style-type: none"> Delay review of information/data access needs (may affect timing; likely to need negotiating or disclosure agreements pre-start up)
<ul style="list-style-type: none"> Agree who manages all aspects of sign-off/commissioning and (if different) who project manages (including external contractors) 	<ul style="list-style-type: none"> Neglect the need to prioritise any internal roles (including project management) if the evaluation is to deliver on time
<ul style="list-style-type: none"> Agree focus of how much method guidance to give to contractors pre-commissioning (and who answers queries) 	<ul style="list-style-type: none"> Forget credible findings may need independent analysis or validation (this may need an external role to be added and this will affect resourcing and timing)

<ul style="list-style-type: none"> • Establish needs for any formal steering or progress review (what for, when and who) 	<ul style="list-style-type: none"> • Forget to brief those recruited to steering on goals and agenda; their roles and agreeing any 'rules of engagement'
<ul style="list-style-type: none"> • If internal evaluation: Identify who fills what roles for direction; design; delivery/data collection; analysis/verification; reporting 	<ul style="list-style-type: none"> • Ignore the need for an evaluation champion – who will have the role of taking forward the evaluation findings and advocating change against the evidence with decision makers

2. Outcomes needed of the evaluation	
DO ...	DO NOT ...
<ul style="list-style-type: none"> Critically review your overarching aim statement for the evaluation – is it clear, credible and realistic? 	<ul style="list-style-type: none"> Defer seeking wider agreement on the aims and objectives; aims must precede decisions on design and are not retrofitted
<ul style="list-style-type: none"> Critically review the subsidiary objectives – are they consistent with the rationale for what is being evaluated? Is it coherent with any logic chain/theory of change for the ‘intervention’? 	<ul style="list-style-type: none"> Forget to use objectives to clarify/set out the specific areas where evaluation evidence is needed to aid decision-making
<ul style="list-style-type: none"> Critically review the coverage of the objectives – do they un-necessarily overlap or duplicate each other? If so, consolidate. Establish any gaps in aims/objectives; is anything missing. How do the objectives change to reflect any gaps? 	<ul style="list-style-type: none"> Confuse objectives (what/how evidence is to be used) with ‘method’ (i.e., how to get evidence); method guidance or prescription follows objective setting
<ul style="list-style-type: none"> Assess realism of aims and objectives; the goals of the evaluation need to reflect the context, time and resources available 	<ul style="list-style-type: none"> Hold back from asking for clarification or challenge – setting solid and appropriate expectations are the foundation of effective and usable evaluation
<ul style="list-style-type: none"> Assess viability of aims and objectives; are they consistent with likely information availability, existing data or other evidence which is ‘to hand’ or can be gathered? 	<ul style="list-style-type: none"> Extend the aspiration for the evaluation beyond the needs of the aims and objectives; information and evidence is a tool and not just ‘nice to know’

3. Timing and delivery	
DO ...	DO NOT ...
<ul style="list-style-type: none"> • Take account of 'upstream' needs (e.g., internal and/or external sign-off of specification); procurement notice period; marketing/tendering/commissioning decision-making lags; etc. 	<ul style="list-style-type: none"> • Forget to allow enough time also for potential contractors to produce viable tenders (2-4/5 weeks depending on needs)
<ul style="list-style-type: none"> • Build in 'engagement time' to liaise with stakeholders (i.e., specification/pre-start-up; during evaluation/steering; pre-reporting including previews of findings; review and sign-off of reports) 	<ul style="list-style-type: none"> • Assume stakeholders are best held at 'arms-length' until findings are finalised; earlier engagement brings challenges/delays but can help later with the credibility of findings
<ul style="list-style-type: none"> • Allow appropriate time for sensible measurement of outcomes (and impacts) – these may take time to be realised; compressed timeframes may miss/under-represent achievements 	<ul style="list-style-type: none"> • Skimp on time for design, testing and clearance of evaluation 'tools'; rushed design compromises information quality and reliability
<ul style="list-style-type: none"> • Allow sufficient time for gathering any new/additional evidence (e.g., survey response/reminder time) and thorough analysis and interpretation by evaluators 	<ul style="list-style-type: none"> • Forget 'good' evaluators will need time for verification of the evidence they do collect; verification also adds to quality and credibility
<ul style="list-style-type: none"> • Build in time for staged/mid-point review (e.g., via contract review or steering group); this is especially important for formative evaluations 	<ul style="list-style-type: none"> • Under-estimate the amount of time needed for staged review within evaluations (especially where steering groups are involved)
<ul style="list-style-type: none"> • Allow for 'downstream' time after (draft) reporting to review, reflect on (consult if appropriate) and sign-off the evaluation before getting results/implications to decision makers, etc. 	<ul style="list-style-type: none"> • Under-estimate time needed downstream to build credibility, confidence and understanding of findings among intermediaries, stakeholders/doubters; evaluation utility may depend on this

4. Use and users of the evaluation	
DO ...	DO NOT ...
<ul style="list-style-type: none"> Focus the evaluation approach, scope, timing and communication on the primary user(s). This will have been agreed from 'roles and responsibilities'; BUT ... 	<ul style="list-style-type: none"> Forget the secondary users ... appropriate engagement will help build the credibility and utility of the findings; are there other (non-user) stakeholders who also need to be engaged
<ul style="list-style-type: none"> Clarify pre-specification how the evaluation findings are to be used; are there any expectations of change/improvement, etc. 	<ul style="list-style-type: none"> Forget that different users (primary and secondary) may have different expectations of the evaluation and its utility; unrealistic expectations of change need to be countered/conditioned for all
<ul style="list-style-type: none"> Identify critical timings/decision making points and align scope and approach to meet these (where appropriate) 	<ul style="list-style-type: none"> Forget that compressing the approach/scope to meet decision-making schedules may mean compromises need to be agreed, with evaluation aims/objectives re-engineered as appropriate
<ul style="list-style-type: none"> Identify if there are critical 'user' intermediaries (people, functions or bodies between whoever is accountable for the evaluation and decision-makers) 	<ul style="list-style-type: none"> Underestimate the importance of champions/brokers of the evaluation findings (positive and negative) in influencing change; findings rarely speak for themselves among decision-makers
<ul style="list-style-type: none"> Identify sufficiently early if/what communication strategy is needed to bring findings/implications to the various users 	

5. Resourcing the evaluation appropriately	
DO ...	DO NOT ...
<ul style="list-style-type: none"> Recognise that resources are your budget, staff and time; these will vary with internal or external evaluation 	<ul style="list-style-type: none"> Underestimate the staff resource and range of skills needed for internal evaluation; external advice or peer review may help build your confidence where the skills mix/experience is limited
<ul style="list-style-type: none"> Remember that 'appropriate' resourcing is led by the scope, needs and expectations of the evaluation – not availability of budget/time, etc. Limited resources may need compromises to the aims and scope 	<ul style="list-style-type: none"> Be funding-led (what can we do for the money); critically review if the budget available is appropriate for the aims and objectives (and/or proposed approach/scope)
<ul style="list-style-type: none"> Appropriately resource project/contract management; this takes time to do well. Does the allocated staff member have the necessary availability, skills and experience? 	<ul style="list-style-type: none"> Forget that project managers will need to balance the added demands of evaluation management with their other tasks/roles; does the new role have clear prioritisation/sign off?
<ul style="list-style-type: none"> Are internal or partner interests/functions 'bought in' to resourcing decisions (e.g., is procurement able to support the necessary timetable) 	
<ul style="list-style-type: none"> Set up appropriate review/steering arrangements pre-evaluation with clear briefing on roles/responsibilities to ensure engagement and continuity across evaluation 	
<ul style="list-style-type: none"> Ensure timing challenges are reflected in the agreed timetable 	

Resource E Making evaluation designs proportional

Influences and considerations in different evaluation circumstances

Making 'proportional' choices in evaluation design will involve balancing the nature and context of the intervention with decision-making needs.

Influence or context	Considerations in evaluation choices
High-profile interventions	High-profile responsible gambling interventions may be expected to have a transformative or high impact. They are likely to require thorough and probably large-scale evaluation evidence, conducted independently, to build a sufficient evidence base to demonstrate effectiveness, value and sustainability and transferability (to other contexts).
High level of innovation in intervention	Highly innovative interventions, even at pilot scale, are likely to require very high standards of evidence robustness (analysis and demonstrable validity) to show how well they are working in practice, their impacts and the scale of returns/added-value. High innovation may also be high risk (for the funder or implementers) which will place added demands on the transparency of the evaluation and how it is conducted and reported.
Short duration of intervention (or required evaluation)	Fast turn-around and intensive interventions provide more limited opportunities for data capture, comparison analysis over time or for longitudinal review. This will need to be reflected in the scale and depth of the evidence collection and review. Interventions which only have the need for an evaluation added part the way through will also have constraints on design, and the likely reliability of evidence.
Large-scale or complex interventions	Larger-scale responsible gambling interventions are those with substantial investment which will probably require a more extensive and engaged evaluation to demonstrate the money is well spent. They also provide opportunities for formative designs or longitudinal analysis and comparisons which need to be exploited if the evaluation is to optimise its usefulness.
Small-scale or pilot/trial interventions	Pilot, trial or other small-scale interventions will have proportionality influenced more by the evidence needs and requirement for fairly immediate decision-making to inform repetition, scale-up or roll out.
Need for wide generalisability of evidence from the interventions	Specific interventions may also be looked to for wider lessons or transferability; an extensive evidence base may be needed to ensure that the results can be generalised with greater confidence.

<p>High level of complexity of interventions (e.g. multiple inter-related activities)</p>	<p>Multi-faceted interventions (e.g., multiple activities) or those addressing a variety of 'needs' or user/beneficiary circumstances mean effectiveness or effects may be more difficult to isolate. Complexity means proportionate approaches are more likely to need to be extensive and sensitive to different user, social or geographical groups or application contexts.</p>
<p>Weak (no) pre-existing evidence base for interventions</p>	<p>Where the existing evidence base is poor, baselines are lacking or comparative evidence is thin, an evaluation is likely to require more extensive evidence-gathering to fill these gaps and to retro-fit 'benchmarks' to contrast how well the intervention performs.</p>

Resource F Internal vs external evaluation – making the right choice

Making a decision about conducting an evaluation ‘in house’ (internal) or commissioning it externally is about more than budget or available resources. Making the right decision here needs to balance up many pro’s and con’s. It will also set the parameters for what can be done, how it can be done ethically, and the quality and credibility of the evidence and findings.

Type of evaluation	What and who is involved	Pro’s	Con’s
Internal or ‘self’ evaluation	An evaluation which is largely (or wholly) designed, delivered, analysed and reported by the organisation which is also delivering or funding whatever is being evaluated.	<ul style="list-style-type: none"> • Cheaper (although staff costs may be hidden or ‘lost’ in the delivery budget) • Less likely to suffer data-access constraints (e.g., Data Protection Act) since much of what is needed for evidence gathering (or access to participants) will be accessible in-house • Making use of ‘contextual knowledge’ with the evaluator likely to be close to the intervention, its delivery and policy context • Quicker, providing for a (usually) faster start (possibly vital for intensive evaluations) 	<ul style="list-style-type: none"> • Evaluation staff may not always be ‘hand-picked’ for their knowledge and evaluation experience (risking quality and credibility) • Evaluation staff may have limited skills and understanding of different evidence collection, analytical, validity and reliability testing options • Internal resourcing priorities may intervene, with the evaluation deprioritised, delaying findings • Evaluation staff may be regarded with suspicion by participants or customers, and there may be circumstances where it is unethical for operators to engage with vulnerable people such as problem gamblers or those at risk. Staff may be wary of sharing their views honestly with internal evaluators.

			<ul style="list-style-type: none"> • Stakeholders may see internal evaluators as 'conflicted' • Lacking independence and (perceived) impartiality with consequences for reduced confidence and credibility in findings
External evaluation	Commissioned by the delivery organisation or direct funding body from outside, using non-aligned evaluators who may be consultants, academics or from independent research groups.	<ul style="list-style-type: none"> • Recognisably independent evaluators, with no conflicts of interests (where screened out in procurement) • 'Expert' (if well selected) with access to specialist expertise/techniques which may provide for more robust and/or cost-effective evidence collection and analysis • Able to draw on wider comparative knowledge, perhaps about similar policy initiatives or interventions which may be of additional value to the evaluation analysis • External evaluators are more likely to be viewed with less suspicion by participants or customers who may be more open and honest about their experiences • Impartial and likely to provide for more confidence and credibility in 	<ul style="list-style-type: none"> • Likely to require formal procurement; may cause delays to start (and end) of evaluation and need a longer overall timetable • Requires an allocated budget and will cost more than internal evaluation • May encounter data access constraints such as data protection limitations to personalised data which could be difficult to overcome • External evaluators may be regarded with more suspicion by delivery staff (although not by participants or customers)

		findings and be taken more seriously by third parties	
Hybrid evaluation	Evaluation is designed, delivered, and analysed by the delivery or funding organisation (i.e., internal) but an external and independent 'peer reviewer' is appointed to critically review method and/or findings – with their review taken into account by, and potentially published alongside, the final evaluation.	<ul style="list-style-type: none"> • Retains benefits of internal evaluation (cheaper, faster, few if any data-access constraints, and using 'close-to-intervention' contextual knowledge) • Introduces an external, independent and impartial validation (if peer reviewer is appropriately selected and not conflicted) • Draws in professional expertise and wider knowledge (although this may be after the event) 	<ul style="list-style-type: none"> • Does not overcome all 'internal evaluation' constraints and the process will still be at risk from changing internal priorities and delays to findings • Likely to require a (smaller) contracting budget and formal procurement to obtain an expert and impartial peer reviewer • Peer review process will result in some delays to the conclusion of the evaluation • Peer review process risks raising method weaknesses in internal evaluation and/or a contradictory assessment of findings which will need to be managed with third parties

Resource G Setting the tone for using evaluation findings

Findings do not speak for themselves. A well-designed evaluation with a robust report does not guarantee its findings will be valued, understood and used. Getting findings into action also depends on how you plan and prepare for this, getting the right stakeholders engaged early and throughout, and how imaginative you are in communicating – not just reporting – before, during and after the evaluation.

Before the evaluation starts ...	
Pre-launch press release	<ul style="list-style-type: none"> • Kick start press release or briefing announcing evaluation, scope, intent, evaluators and engagement process, etc.
Start-up briefings to key stakeholders (and others likely to influence the take-up of findings)	<ul style="list-style-type: none"> • Selected 1 to 1 meetings/discussion to key third parties • Engagement workshop or event to raise awareness and inform expectations
Pre-launch calls for evidence	<ul style="list-style-type: none"> • Supplementary to mainstream (systematic) evidence collection via the evaluation
During the evaluation ...	
External appointments to evaluation steering or advisory group	<ul style="list-style-type: none"> • Balance 'internal' appointments to evaluation steering/advisory group with external members representing third party interests
Staged or interim release of early or mid-point findings	<ul style="list-style-type: none"> • Staged release of findings to provide early sharing of evidence • Usually prior to preliminary conclusions
Invitations to comment on early findings	<ul style="list-style-type: none"> • Drip feeding findings to invite wider comment and engagement in the evaluation
Open/limited invitation to comment on draft final report/findings	<ul style="list-style-type: none"> • Prior to final reporting • Selected or open invitation to discuss and review draft conclusions
In parallel with final reporting ...	
Written press releases and briefings	<ul style="list-style-type: none"> • General and targeted press releases • Pre-final report release briefings
Policy/practice briefs, cascade briefing events	<ul style="list-style-type: none"> • Tailored 'punchy' briefs to specific or priority audiences and/or third parties

Own and stakeholder communication channels	<ul style="list-style-type: none"> • Own media and also harnessing stakeholder media, etc. • Customise existing communications material for users/stakeholders
Presentations at events, workshops, conferences, etc.	<ul style="list-style-type: none"> • Launch event/conference or stakeholder workshop • Roll-out 'evidence and where next' events • Priority interest group briefings or seminars
Blogs, podcasts, e-fora	<ul style="list-style-type: none"> • Harnessed to re-enforce, widen and maintain engagement and communications momentum
Twitter and misc. social media	<ul style="list-style-type: none"> • Wide distribution of highly condensed key fact or 'single messages'
Practice toolkits, guidelines	<ul style="list-style-type: none"> • Specialist evidence-based guidance to wider users (e.g., better practice)

Some useful sources of further information and guidance

General guidance

CES (2008), *Practical Monitoring and Evaluation: A Guidebook for Voluntary Organisations*. 3rd edition, National Council for Voluntary Organisations, London

See also from NCVO guidance resources for evaluation at:

<https://knowhownonprofit.org/organisation/impact>

Her Majesties Treasury (2011), *The Magenta Book: Guidance for evaluation*. Magenta-Combined, London:

www.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

United Kingdom Evaluation Society (2013), *Guidelines for Good Practice in Evaluation*. London: UKES

www.evaluation.org.uk/images/ukesdocs/UKES_Guidelines_for_Good_Practice_January_2013.pdf

Useful further sources and reading

Fox, C. Grimm, R and Rute, C. (2016), *An Introduction to Evaluation*. London: Sage

Hall, I and Hall, D (2004), *Evaluation and Social Research: Introducing Small-scale Practice*. Palgrave Macmillan. Basingstoke

Mertens, D.M and Wilson, A.T. (2012) *Programme Evaluation Theory and Practice: A Comprehensive Guide*. New York. Guildford Press.

Parsons D (2017), *Demystifying Evaluation: Practical approaches for researchers and users*. Policy Press. Bristol