Measure What Matters Guide

Contents

Introduction

One of the key principles of Happy Museum is to 'measure what matters'. We want more than just the numbers of people coming through the door; we want to understand individuals' experience and the impact of museums on emotions, creativity, wellbeing and sustainability.

We have compiled a collection of tools to support you to measure what matters to you and the individuals you work with. These tools help you to plan your work, gather evidence of your impact, and understand the difference your organisation makes. Measuring what matters will help you to improve and grow as well as to tell the story of your museum to others.

Each tool is different and has its own strengths. Some tools are very simple and quick, others are more complex and require more time and planning. The general rule is the greater the demands of the tool, the more meaningful, robust, and reliable your results will be; you get out what you put in. We know people are stretched and resources are under strain so the tools you use will likely be a trade-off between what you want and what is possible.

The descriptions below will help you identify the tool or tools that meet your needs, suit your project, and match the resources you have available. However much time and resources you have, or whatever the characteristics of the group you are working with there is a tool here that will help you to measure what matters.

Measuring what matters is a journey and getting started can be the hardest part. Beginning to collect simple data will put you in a stronger position than if you have no evidence at all and you will learn as soon as you start what works well and where there is room to improve. You will be able to hone and build on the methods you use.

Sometimes you will want to get evidence from the same people before and after a project (longitudinal research) but sometimes it will be enough to get evidence from groups of people which might be broadly, but not exactly, the same. The best option is to use more than one tool in combination to provide a richer picture of change and to understand the experience of your participants through different lenses. Be clear and transparent about which approaches you are using.

There are two key points to keep in mind:

- We need to be able to understand the situation before and after the project so we can see what has changed
- We need to account for what would have happened anyway. Ideally, this should be done by monitoring the changes in a control group. This means collecting data from a group of people who are not involved in your project (as well as from those who are) to give you a point of comparison. All tools can produce more meaningful results through the introduction of a control group. If this isn't possible, we have to estimate what would have happened anyway (known as deadweight).

The tools will support you with key stages of measuring what matters:

1 **Planning and review for organisational change**: First, we need to establish what we are trying to measure by agreeing what we expect to change, what the indicators would be, and how the information will be collected. We recommend using Story of Change for this but there are other tools than can help establish the context and need of your project.

2 **Baseline and review – collecting evidence of change**: We collect evidence of change to find out if our Story of Change is sound and if the links we have made between what happens and the difference it makes stand up, as well as quantifying these changes.

3 **Review for personal or organisational change**: Sometimes it's not possible to collect before and after data, but evidence collected at the end is still valuable.

4 **Analysing evidence**: We need to analyse the evidence and ensure the results are statistically significant, in the case of Happy Museum this has been ensured by involvement of the 'happiness economist' Daniel Fujiwara.

5 **Learning what made the change**: Having ensured our research is valid, we go back to the Story of Change to flesh out our learning and clarify what made the change. We update and improve the Story of Change, and identify what further research is now necessary, either because there are negative results we need to understand, or we want to know more, or things have changed. So we create a cycle of continuous learning and improvement.

With any tool that you use, you will need to explain to your participants why you are asking for their input and contribution. This text might be useful to communicate the context:

"Our museum is a Happy Museum. We think Happy Museums are more enjoyable, sociable and interesting places than traditional museums. To find out if this is true, we need your help with information so that we and others can improve what museums do"

Planning and review for organisational change

Story of Change (SoC)

A planning and review tool

- video
- Story of Change planning doc
- Story of Change template.

Story of Change is a process to help structure your project, to support you to focus on the difference you want to make, and how to achieve it. The SoC process (which is similar to Theory of Change or logic modelling) will help clarify your overall purpose, the individual changes you hope to make, the steps to get there, and the links between those steps. It is often tempting to focus on what you do rather than what you want to achieve. SoC will help maintain focus on what you set out to achieve and ensure you consider everything that is part of that.

Story of Change is a method used for participatory evaluation. It is the first step in establishing a Social Return On Investment (SROI) analysis, an approach that is all about measuring what matters. Story of Change has been used by every commissioned Happy Museum. Most have found it very useful, and for some it's changed the way they work. For example IWM North used it to plan their whole-museum development, and the LTM learning team has now embedded its use in their programme planning for family, community, young people and schools services. It looks like this:

Planning what we want to achieve

Who?	How?	What?	Why?	Why?
The people who matter	Drivers and investment	Delivery	The difference we make	A vision

Delivering what works

Who?	How?	What?	Why?	Why?
The people who matter	Drivers and investment	Delivery	The difference we make	A vision

When should we use it?

SoC is valuable in both the planning and evaluation stages. Using SoC to plan your project will help you put your vision at the heart of your work, to understand the resources you need, your activities, your milestones, and to be aware of risks and challenges. SoC also provides the basis of a meaningful evaluation as it will outline where you expect to see change and what you should collect to gather evidence of that change.

How do we use it?

A brief overview is provided below. More detailed guidance and a template for SoC are available on our website

The best way to develop your SoC is to use a large piece of paper and different coloured post-it notes.

1. Establish your overall vision - what is it that you hope to achieve? Your project will work towards this vision and other things may also be necessary to make it happen.
2. Agree the people who matter - the people who either experience the change or are important to make it happen. Think about knock on effects to other stakeholder groups as well as the immediate participants.
3. Clarify the difference the project will make – why are you doing it?
4. Check these are not actions but are changes or outcomes, for example, changes in understanding, behaviour, knowledge, attitudes or situations.
5. Agree the activities you need to deliver results and be explicit about who is involved.
6. Clarify the drivers or inputs necessary. This might be resources, knowledge, good practice etc. – everything you need to deliver the activities
7. Key questions to reflect on your SoC – Does the logic stack up? Do activities lead to desired results? Do these result in outcomes that will lead to your visionary change? If you do what you've planned will you get where you want to be? Is it achievable?

Usable versus robust?

How useful and empowering?

SoC can be really useful to ensure you understand your project, informing design and delivery, to know if it is achieving what you expect it to, and to feedback findings into improving your project.

SoC is about establishing a thorough understanding of your project, so it can be a demanding process that involves a number of people. However, going through this process is part of the strength of the tool, as it encourages questions to be asked and assumptions challenged. Structuring the conversation well is important to ensure the views of a group of people are all considered and different elements covered. This can be achieved through following the associated guidance.

How can robustness be maximised?

Robustness is about gaining a comprehensive understanding of your project and expectations for what works, a first step in thinking about cause and effect.

The key thing to ensure SoC is done well is to make sure relevant issues and areas are considered carefully by the correct people following every step of the guidance. Consider involving partners, funders, board members, delivery personnel, participants and so on. It is good practice to revisit the SoC on a regular basis to reflect, review and update as necessary.

What can the results tell you and what do they not tell you?

The SoC process helps you to structure your thinking about the delivery, monitoring and measuring, and review of your project. It helps you to discuss what works with others. It is a case of getting out what you put in. It will not be useful if you don't put the initial effort in to create a meaningful SoC. Your SoC should be an active document that sits at the heart of your project, is regularly revisited, and informs what you do.

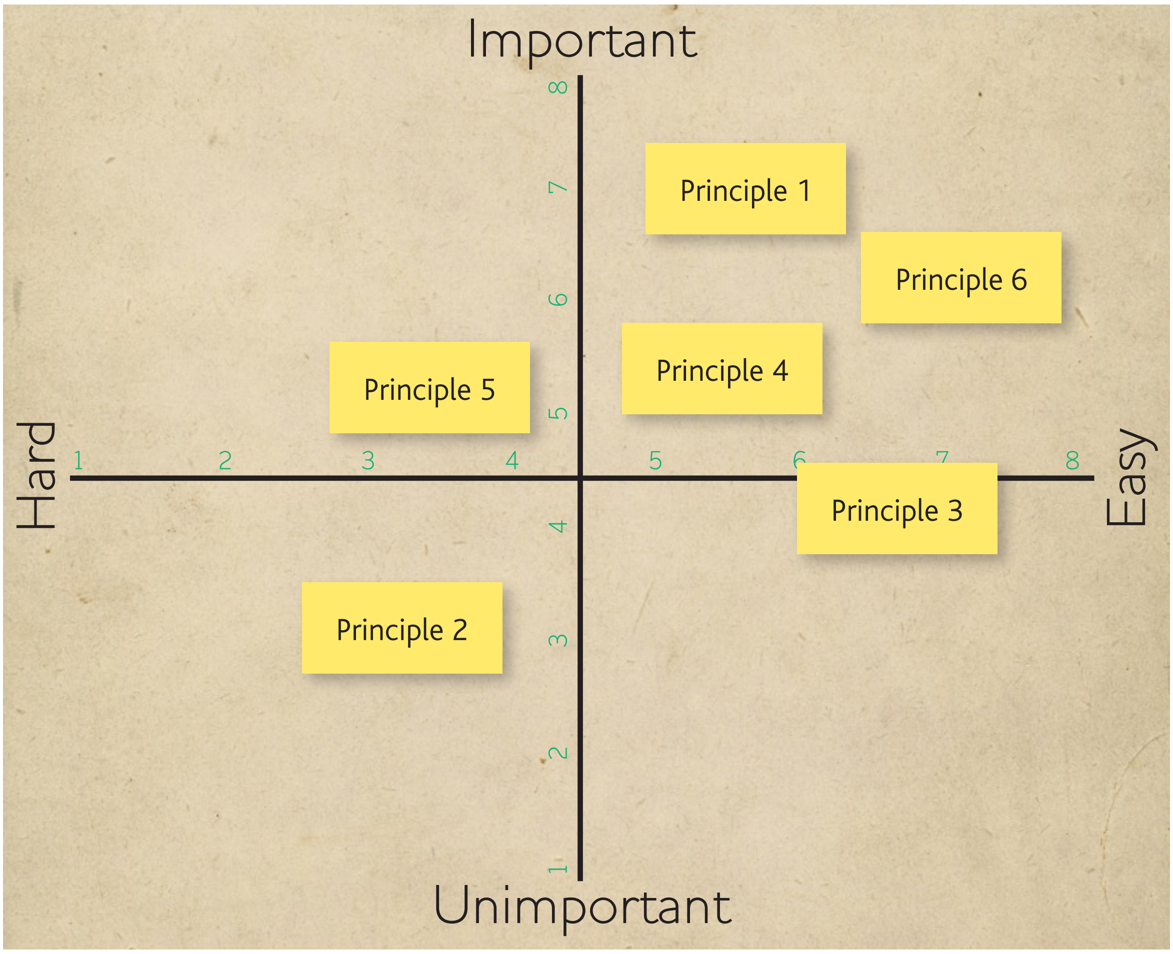
Behaviour Attitude Matrix – BAM card sort

A planning and review tool

- BAM card sort

This tool can inform the prioritisation of different project elements, such as the Happy Museum principles, using input from participants based on a 2x2 matrix. Done before and after it gives feedback on how these change after the project. Participants consider each element against two questions that can be defined by the project managers, for example, 'important' (attitude) and 'easy' (behaviour), or 'fun' and 'good for you'.

The BAM card sort has been used by every commissioned Happy Museum to give feedback to the overall programme. We recommend it because it prompts very useful conversations and provides a simple snapshot of how people are responding to Happy Museum over time. It looks like this:



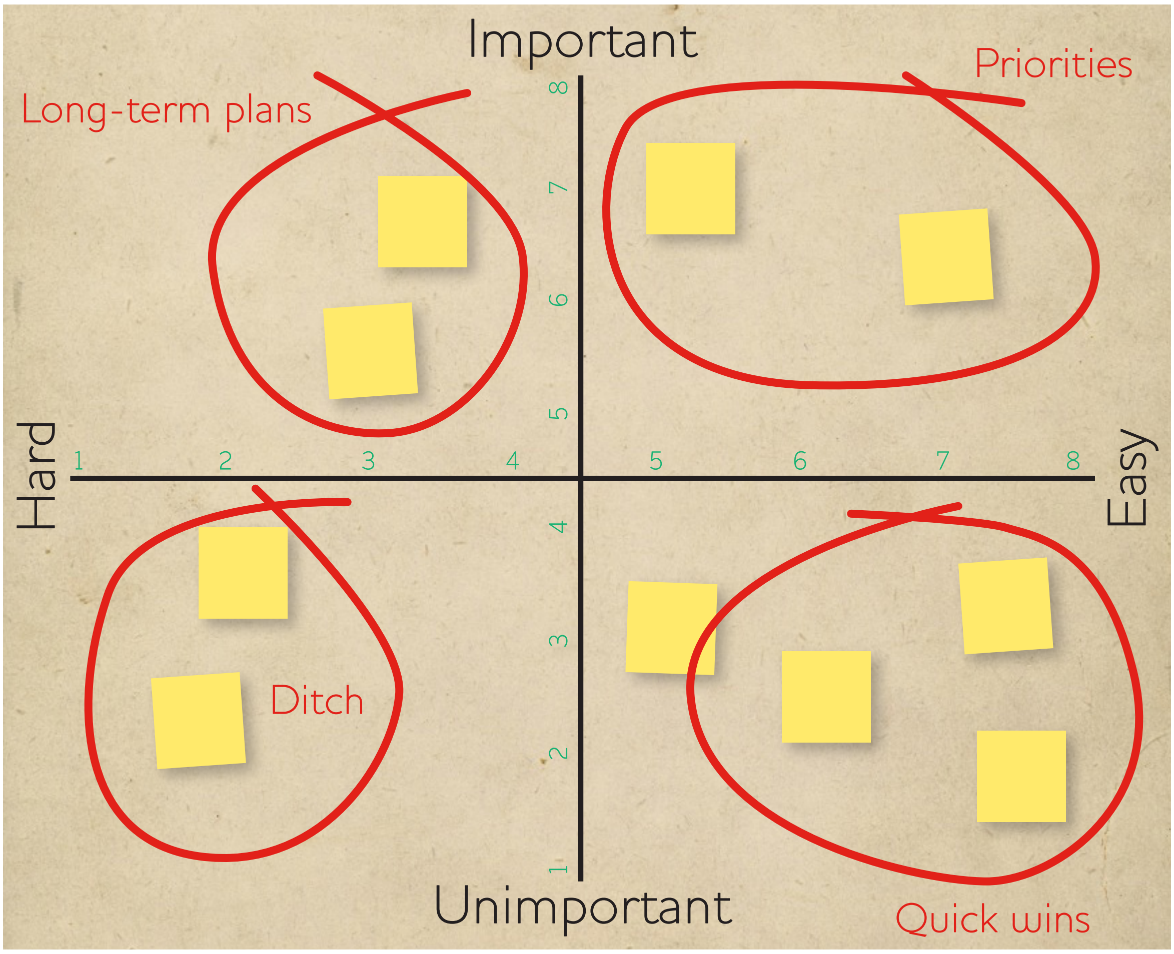
When should we use it?

We have used BAM to compare behaviour with attitude and change over time, but the BAM is also useful when you need to make decisions around what to deliver with the limited resources you have available. It's likely you'll have different objectives to deliver against and need to balance activities across the two, for example, you may want to deliver a project that is fun (in order to engage people and keep them engaged) but also something that is good for people (to help improve their lives). Different project activities are likely to have more of a focus on either being fun or being good for them.

You can use this tool in the planning stages although it may be easier to use at the reviewing stage once people have had experience of different elements. Often in change programmes things get worse before they get better. (See our Transformation tool) For example various principles might be judged easy until participants have really understood what they entail. So running the card sort before, after and beyond the project can help uncover those sorts of experiences.

How do we use it?

- We want to consider each project element/activity against two questions that can be defined by the project managers, for example, 'how easy is it for us to do?' and 'how important is it?'
- Draw up a chart on a large piece of paper with these questions on the axes (see example below)



- Ask participants to place post-its/card onto the chart on one axis first, for example, 'how important is this?'
- Ask participants to move the cards along on the other axis, for example, 'how easy is this?'
- The position of the different elements then reveal what your priorities should be (the ones that deliver on both objectives in the top right hand corner)
- You can then repeat the process to see if you can change how an element performs in the matrix, for example, can you make something that is very good for people more fun.

Usable versus robust?

How useful and empowering?

Once participants understand the concept of the matrix, the BAM is an easy and friendly way to identify how different elements deliver on project objectives and whether people have learnt or changed their attitude. Participants will need to see that their contribution has been acted on. Depending on the environment in which is it done certain biases can come into play that are worth being aware of. Participants may feel they should report something positively as they feel they should be grateful that they were involved, if the end of an experience was particularly pleasurable they may misremember the experience of the entire duration of the activity.

This is a simple tool that can identify the suitability of different project elements, like the Happy Museum principles on the delivery of objectives. The position of the post-its should inform decisions in the design and delivery of an effective project. It can give an indication (though not proof) of whether awareness of an issue has grown as it becomes harder, or whether something has become easier after the project but converting positions into numeric scores.

How can robustness be maximised?

- If you are using the BAM to measure the change in performance of elements on different dimensions make sure you collect data at the very start of the project so you have a baseline to which you can compare your results
- Provide exactly the same guidance to participants each time and don't say anything that may lead their response
- Remind people they should be honest and open
- Give participants the opportunity to place their post-its anonymously
- Always conduct the exercise at the same time of day and on the same day of the week
- Because the matrix doesn't show a score and participants aren't reminded of their original feedback, we stand the best chance of getting an biased answers

What can the results tell you and what do they not tell you?

Results can reveal the performance of specific elements on two dimensions from the participants' perspective (being aware of aforementioned biases). You can only ever compare the different elements on two dimensions at one time, which can be restrictive. The results cannot prove that your project impacted behaviour or attitude changes, but they can give an indication that change has taken place.

Tipping Point

A planning and review tool

- Tipping point

When planning your project, identify what you think will make the biggest difference in the organisation; what could create a 'tipping point' in organisational change according to the culture and practices of the organisation (as opposed to with your audiences). This is best achieved through engaging a diagonal slice team, which is a group of individuals at different levels and from different teams in the organisation or partner organisations.Tipping Point was introduced at a Happy Museum symposium.

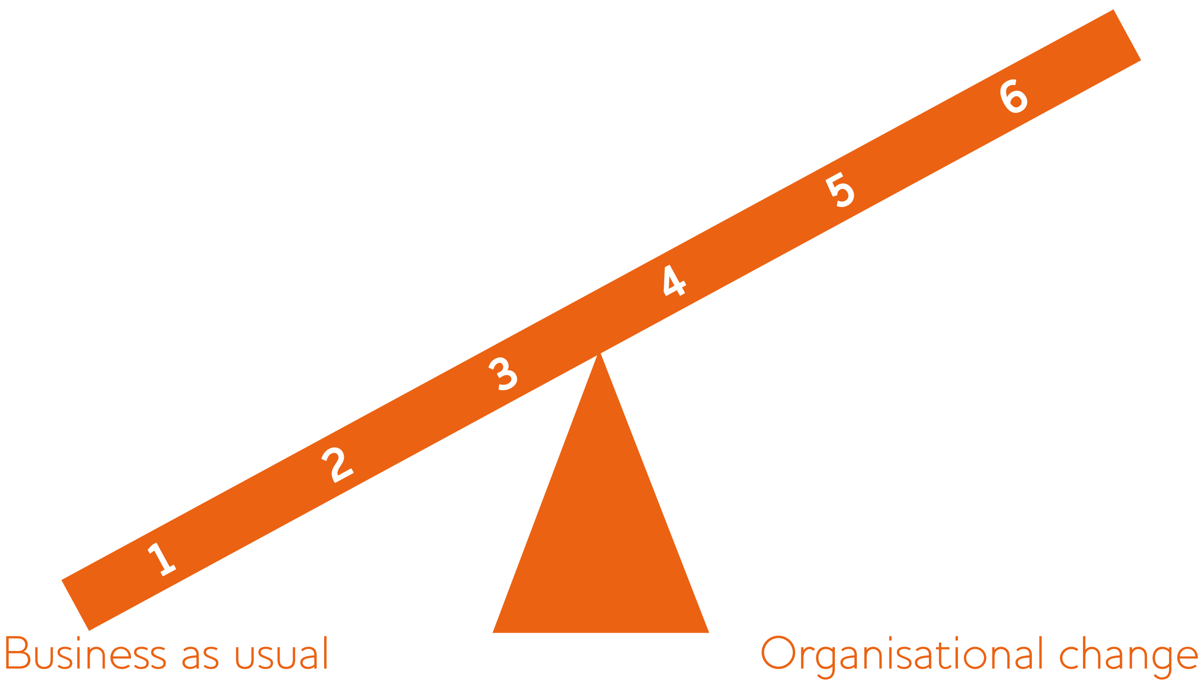
This can inform the approaches to make important things happen and the decisions about how to collect evidence to show that you've kept to your plans.

When should we use it?

It makes sense to do this either after your Story of Change exercise to identify the key factors, or you may want to do this first to inform your Story of Change.

How do we use it?

- You could use the cards from Happy Museum guidance or create your own themes. Your own themes could simply be ideas you feel could make a considerable difference your organisation.
- Place the cards on the scale diagram. Position things on the right that you think will make the most profound difference, and on the left the things that will make little difference. Those in the middle you might be unsure of, or think are only of medium importance. The scale is just a visual aid to help staff focus on where the balance of power lies, and what might tip the organisation into change. Most of the value will come from discussing these things together, so don't avoid getting together what might be a challenging group.



- Re-visit the process after the project to agree what did make the biggest difference, and whether anything has created a tipping point so when you need to make future organisational developments you have a better idea of how to generate change. Did anything unexpected make a big difference?

Usable and friendly vs robust?

How useful?

Tipping point is an introductory tool, useful to facilitate a discussion between people who may not normally work together, to pull out the factors that have an important role in organisational change, and to then enable a focus on what's important to the organisation as a whole.

Tipping point should be empowering to those involved especially to more junior people as their voice should be considered as equal to those at a senior level.

Tipping point may lead to some lively debate as it is likely that people from across the organisation will have different views about what is important. This shouldn't be avoided as it should lead to useful insight. It is a very usable and useful tool. The tool is not designed to be robust for evaluation, but it will nonetheless help you review afterwards what did make the difference.

How can robustness be maximised?

- Set up a representative, diagonal slice team
- Decide on a facilitator who will make sure everyone's voice is heard and given equal weight
- Define rules for everyone to agree to in the discussion, for example, everyone should be honest and open, everyone gets the opportunity to speak, everyone should actively listen, the outcome of the discussion will be written up and agreed.
- Keep an accurate record of the results
- When you come to review, try to include the same facilitator and staff in the exercise.

What can the results tell you and what do they not tell you?

The results can identify what is important in leading to organisational change. The results do not tell you if you have managed to identify the key factors, or tell you how to go about making sure they happen. It is one piece of the jigsaw.

Transformation Cycle

A planning and review tool

- Transformation cycle

This approach is for organisations that are used to thinking about how they improve. It is a more sophisticated framework for planning and reviewing change, to the point of transforming the organisation by really embedding new practice. It also recognises the importance in Happy Museum work of the active individual in an organisation,by combining learning from Maslow about behaviour change with learning from Kotter and others about organisational change. Again it is for use with a diagonal slice team.

When do we use it?

It makes sense to do this either after your Story of Change exercise to identify the key factors, or you may want to do this first to inform your Story of Change. It might be an approach to repeat regularly for monitoring progress.

How do we use it?

You could use Happy Museum learning or your own thinking, testing each element against the transformation 'maturity' cycle. This asks you to establish whether you are in one of four quadrants of a cycle:

1 'unconscious incompetence' where there is not yet a motivation for change

2 'Conscious incompetence' where the need for change is being acted upon and old practices are beginning to be 'unfrozen'

3 'Conscious competence' where new practices are put in place but still require attention

4 'Unconscious competence' where new practice is 'frozen' in place and becomes business as usual.

The questions prompt you towards activities or milestones that are generally understood to help create successful change.

Usable versus robust?

How usable and empowering?

The tool requires in-depth thinking and at the review stage evidence should be gathered. It is likely to lead to considerable debate, including different perspectives by different teams and staff. As long as the process is followed through it will be empowering by allowing voices not normally heard to influence organisational change.

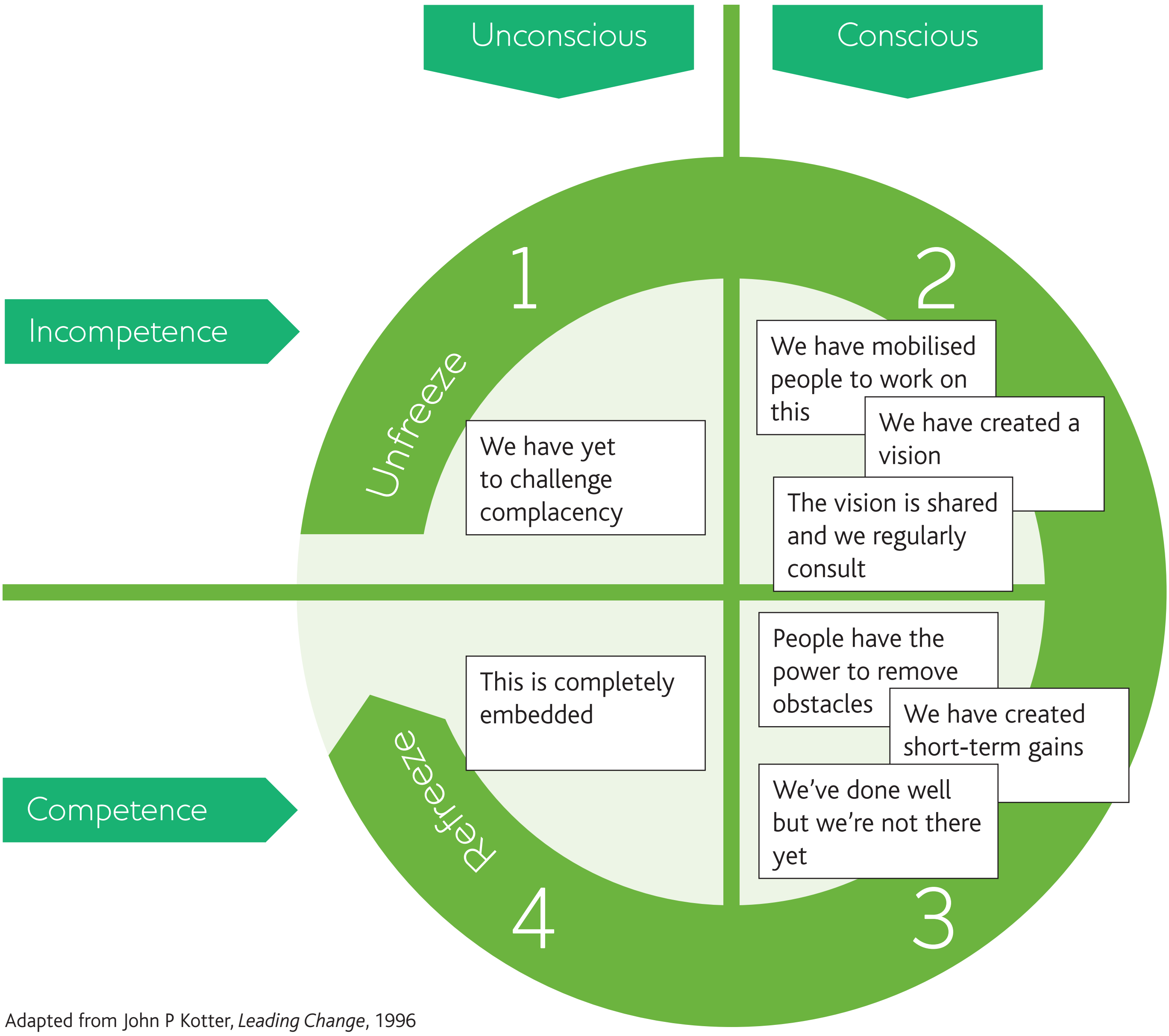
How can robustness be maximised?

- Set up a representative, diagonal slice team
- Decide on a facilitator who will make sure everyone's voice is heard and given equal weight

- Define rules for everyone to agree to in the discussion, for example, everyone should be honest and open, everyone gets the opportunity to speak, everyone should actively listen, the outcome of the discussion will be written up and agreed.
- Keep an accurate record of the results
- When you come to review, try to include the same facilitator and staff in the exercise and ideally some more objective evidence, such as survey results.

What can the results tell you and what do they not tell you?

The results can identify what is important in leading to organisational change and are guided by considerable research into how individuals and organisations change. The results do not tell you if you have managed to identify the key factors, or tell you how to go about making sure they happen. It is one piece of the jigsaw, but by asking the right questions will help to focus on the 'right' things.



Adapted from John P Kotter, *Leading Change*, 1996

Baseline and review for personal change

LIFE Survey

A baseline and review tool

- Full guidance
- Paper surveys
- Survey Monkey
- Self-analysis Template

LIFE is a survey tool developed by Simetrica with Happy Museum and is the most robust tool on sustainable wellbeing, not just in this toolkit, but currently available in the arts and cultural sector. LIFE enables you to assess the impact of different types of museum involvement on a set of positive outcomes within four categories: Learning, Interaction, Feelings and Emotions, and Environmental awareness.

LIFE is versatile as you can use it to measure the impact of the entire museum, a specific exhibition, or a project and on four different groups; volunteers, audience, participants in a project and staff. LIFE also involves surveying people who have not experienced the museum (or surveying these groups when they are not involved with a museum), creating a control group, whose results are used as a point of comparison to the results of those people who are involved. This provides a good measure of what would have happened anyway and so identifies change caused by involvement with the museum.

LIFE was developed by Simetrica, the Oxford Internet Institute and Happy Museum following action research by Happy Museum commissions which used the Five Ways to Wellbeing and other guidance. It was tested with others around the country and turned into an online tool.

When should we use it?

LIFE represents best practice so should be your first choice and used wherever necessary resources are available. It requires more organisation, data collection and input than other tools and so it's important to fully understand what is involved before committing to it. It may be that you are not yet in a position to use LIFE but could be something that you aim towards being able to use.

How do we use it?

See the LIFE guidance on our website for more detailed information on how to conduct the surveys. The surveys are done using an online survey (which you can print for a pen and paper version) with results entered into a pre-prepared spreadsheet. There are then three options. You can use a template to do your own basic statistical analysis, pay for us to run your analysis, or periodically we will run a national analysis.. See links to:

Survey

Template for data input, with simple descriptive analysis

Template for data input, with more complex t-test analysis

Template for data input, with guidance on how to do your own x analysis

Template for project information that allows us and you to understand better your Story of Change

Usable versus robust?

How useful and empowering is it?

LIFE is extremely useful as it can provide robust evidence of change that has happened as a result of your museum around learning, interaction, feelings and emotions, and environmental awareness. It also captures overall life satisfaction which is in effect a catchall that will pick up on any positive impacts in any domain of life whether it be through making friends, feeling healthier, feeling more confident and so on.

Results and Simetrica analysis are of a level of scientific rigour used by central government departments to inform policy decisions as well as international organisations and other sectors in the UK.

Self-analysed LIFE results are less robust, because there is a risk of the data being input with errors and the analysis is basic. (It cannot include the tests which exclude factors beyond the project)

LIFE surveys may empower people as they are consulted in detail on their feelings and opinions. This can give the message that people have a role to play in understanding the impact museums have and knowing that their responses go towards informing key decisions. These feelings of empowerment could be emphasised if results are shared with the survey participants

Results from LIFE surveys are extremely robust. The surveys require organised data collection including collecting data from a group of people not involved with the museum. While it can be an unfamiliar challenge to identify and collect data from a control group, this significantly improves results as we can be more certain that the museum is responsible for the changes we see as other factors impacting on outcomes are filtered out. So while LIFE may take more planning than other tools, you are rewarded with more meaningful and accurate results that are of a standard recognised at government level.

How can robustness be maximised?

- Do the before surveys (if applicable) as close to the very beginning of the project as possible
- Provide the same guidance to all the people administering the survey
- Be consistent with survey mode for example, face-to-face, online, or by phone
- Survey as many people as possible and seek an answer to every question
- Repeat the surveys at the same time of day and day of the week as much as possible
- Ensure control surveys are conducted at the same time as the other surveys

- Check the quality of the data entry i.e. do spot checks to see if the answers from 1 in every 10 surveys has been entered correctly.

What can the results tell you and what do they not tell you?

Results reveal changes in learning, interaction, feelings and emotions, and environmental awareness as a result of being involved with your organisation. Ideally we'd look into a parallel world where the same people were not involved with the project to see what would have happened without. The inclusion of the control group offers the next best practical option to identify the changes you are responsible for.

These results can then be confidently used to inform decisions.

LIFE is a quantitative rather than qualitative tool and so does not reveal any detail about the individual journeys or stories of individuals involved. For this reason, LIFE fits well with other more narrative methods for individuals. Used in conjunction with a Story of Change however, or our project template, you can start to explore qualitatively more about your unique museum offer.

Time Capsule

A baseline and review tool

- Time Capsule

Time Capsule is a relaxed but engaging and fun tool that captures changes in opinion or sentiment connected to an issue or place. This tool was developed by Happy Museum commission Reading Museum and adapted for more general use. Time Capsule could be used to capture the way people feel about a neighbourhood, a museum, or a town centre for example, or it could be adapted to express their personal change using LIFE questions. Participants are invited to select word cards from a collection to represent the way they feel at the start and again at the end of a project to reveal if their feelings have changed.

When should we use it?

The tool can be used whenever a project's aim (or one of its aims) is to change the way people feel towards an issue or area. The exercise should be done at the start and end (or at an agreed point) of a project.

How do we use it?

- Draw up a set of word cards to include a range of descriptive words (positive and negative) that could apply to a particular issue or place such as Strong community. Beautiful. Vice. Avoid. Dangerous. Peaceful. Historic. Modern. Ancient. New. Friendly. Ugly. Safe. Industrial. Crowded. Diverse. Bad. Good. Rich. Poor. Open space. Violent. Good Atmosphere. Unfriendly. Lonely. Cosmopolitan. Decay. Urban. Countryside. Fumes. Children. Elderly. Working class. Youth. Families. Residential. Business. Criminals. Polluted. Dusty. Clean. Dirty. Noisy. Quiet. Bright. Dark at night. Restaurants. Boring. Pubs. Entertainment. Flooded. Alcoholics. Drug users. Gangs. Schools. Medical Centre. Shopping centre. Local market. Police station. Church. Police Cars. Community Centre. Shops. Trees. High rise. Flowers. Brickwork. Concrete. Green. Good Facilities. Good transport. Remote. Arty. Suburban.
- You can categorise these words as 'negative' or 'positive' (plus other categories – see below).
- At the very start of the project, bring participants together and ask them to select words that represent their view of the area or issue and seal them in an envelope or box (you will need to have multiple copies of the cards or you could ask people to copy their selected words on to blank cards).
- It is up to you if you want to seal the words for each individual or for the group as a whole.
- Repeat the exercise at a later date or at the end of the project. Enter the words into separate before and after columns in Excel and sort alphabetically or use the Counta function. This will highlight where selected words are repeated and where they have changed. You can colour-code the negative and positive words to show up the difference

Key	1st evaluation		2nd evaluation		Change
Negative	11	50%	8	32%	-18%
Unique word	4	18%	7	28%	10%
Positive	0	0%	3	12%	12%
Neutral/ unclear	7	32%	7	28%	-4%
Total words	22	100%	25	100%	
	Avoid,		Beautiful,		
	ugly,		historic,		
	crowded,		ugly,		
	diverse,		crowded,		
	poor,		diverse,		
	bad,		poor,		
	violent,		bad,		
	decay,		violent,		
	urban,		decay,		
	working class,		urban,		
	youth,		working class,		
	business,		residential,		
	criminals,		business,		
	polluted,		dirty,		
	dirty,		noisy,		
	noisy,		alcoholics,		
	dark at night,		drug users,		
	alcoholics,		schools,		
	police,		shopping centre,		
	shops,		local markets,		
	drug users,		shops,		
	cars		brickwork,		
			concrete,		
			transport,		
			good		

- Compare the words in each section. In the example above, the 'after' session may see some words appearing less (e.g. Criminals or Dangerous) or see the appearance of some new key words (e.g. Bright or Fun).
- You may want to present your results as the number or percentage of negative and positive words in each session, or use a free online tool such as Wordle to generate a pictorial representation or 'word cloud' of the selected words.

Usable versus robust?

How useful is it?

Time Capsule creates a picture of your participants' view of the area/issue you are interested in and offers an easy-to-achieve but systematic and concise way to get a sense of any changes. The results of Time Capsule are simple to understand and communicate. Opinions and feelings expressed through conversations or focus group discussion can be insightful but difficult to summarise and communicate.

Time Capsule asks participants to consider how they feel about a specific area/issue and so could encourage feelings of connectedness and investment in the project to improve this. The ceremony involved in sealing the responses – or even putting them into the museums store for a period – emphasises how seriously you take this.Alternative methods are to use a set of pictures in the analysis, so it can be used with people with learning difficulties or for whom English is a second language.

The tool is very friendly as it asks the participants to do a very simple and even fun task and requires minimal administration. The tool is vulnerable to many biases (which it is possible to minimise if you are aware of them) as perceptions are likely to be prompted by what is available and restricted by what is not. For example, if a feeling isn't represented by a card then it will be missed, or an available card in the range may not be seen and so not picked even if it is relevant. There is also no consideration of what would have happened anyway without the tool, for example, feelings about an area could be different in the summer compared to the winter regardless of any other factor.

How can robustness be maximised?

- Make sure you do the first exercise at the very start of the project so you have a baseline to compare to
- Be consistent in your approach - provide the same guidance to participants each time, present the same word cards or pictures, lay the cards out in the same order and ask people to begin the exercise from the same position
- The more cards and the broader the range of words you offer, the richer the picture you will be able to capture but if you offer too many you may risk overloading participants. Consider a discussion session before the first selection exercise to inform your decisions around the cards that you should include that will be most relevant to your project and group.
- Consider offering words that work as scales on the same theme, for example, 'boring, interesting, fascinating' or 'dirty, clean' to capture movement.
- Attach tags to words so you are able to categorise results connected to your aims, for example, positive or negative, indoors or outdoors, individual or social, natural and man-made. The first or 'before' exercise may see over half of the words having negative connotations and only a few being positive compared to the second session where a third were positive and a third negative (with the remainder dependant on preference).
- Do the exercise with as many people as possible before and after

What can the results tell you and what do they not tell you?

Results can create a picture of how your participants feel about something, for example their area, and give an indication of how these perceptions change over time, particularly if feelings change between negative and positive. Time Capsule can offer a concise but qualitative record of potentially complex feelings. The approach is not comprehensive as key feelings could be missed and we do not know if the project is responsible for any changes that we observe, as lots of other things will be going on in people's lives.

Where Do You Stand?

A baseline and review tool:

- Where do you stand?

Where do you stand? is a simple tool that can gauge participants' opinions by asking them to stand a line of masking tape to show whether they agree with a statement on a scale of 1-7. It is used frequently in Happy Museum events where people don't know each other.

When should we use it?

Where do you stand? can be great if you're working with a participant group who may be intimidated by form-filling or if you don't have the time or resources to use a more sophisticated approach.

The tool should be used before involvement in your project to capture a baseline and then repeated either at the end of the involvement, or at regular intervals to be able to track changes in levels of wellbeing.

How do we use it?

- Stick a line of masking tape on the floor marked with a 1-7 scale (1 = strongly disagree, 7 =strongly agree).
- Present participants with a statement and ask them to stand on the line to represent whether they agree or disagree with the statement.
- Take a photo and enter the position of individuals into a spreadsheet, or get them to put their initials on the tape.
- To measure change, repeat the exercise with the same individuals and statements at a later date and compare results in a second column in the spreadsheet
- Subtract the before score from the after score to reveal the movement on the scale for each participant on each statement. A positive number shows positive movement, a negative number shows negative movement

	This museum is a good host		
	Before	After	Change
Participant 1	5	6	1
Participant 2	4	6	2
Participant 3	3	6	3
Participant 4	4	5	1
Participant 5	6	5	-1
Participant 6	7	7	0
Participant 7	3	6	3
Participant 8	4	6	2
Participant 9	2	5	3
Participant 10	3	5	2
Participant 11	4	5	1
Participant 12	4	7	3
Average change			1.7

Or

Add up the totals for all individuals before and for after to reveal the movement of the group overall.

You could ask any number of questions. You could have a look at the LIFE surveys for suggestions for other questions you could include. It is also possible to use this in a light touch way, without the tape or recording responses, to get a sense of the feelings of individuals in the room or to open up discussion.

Usable versus robust?

How useful?

The tool requires minimal organisation and record-keeping so it's useful if you don't have the capacity to use other measurement methods. Where do you stand? provides an idea of how the self-reported wellbeing of individuals is changing over time and does not require participants to do any form-filling or interviews, which can be intimidating for some groups.

The tool asks individuals how they feel that demonstrates the museum is interested in them, which can empower individuals. We find it prompts very useful discussion.

Where do you stand? requires little effort is very friendly and easy to use and can be used as an ice-breaker at the start of a project. You don't need any resources other than time and a pen and paper (although it can be improved with extra resources - see below). It is important to be aware that the tool is open to many biases that may affect how people report their wellbeing, for example, feeling they should report higher levels, worrying what other people will think, being influenced by others, or being overly affected by irrelevant circumstances, such as the weather.

How can robustness be maximised?

- Make sure you collect data at the very start of the project so you have a baseline
- Make sure the wording of each statement is available on a slide or flipchart so people don't have to remember it
- Provide the same guidance to participants each time and don't say anything that may lead their response
- Remind people they should be honest and open
- Ask people to do it individually rather than as a group to reduce being influenced by others
- Always conduct the exercise at the same time of day and on the same day of the week

What can the results tell you and what do they not tell you?

The results can provide an indication of the change in wellbeing for individuals whilst they are engaged with your project. It is important to remember the possible biases in how they report as well as the caveat that we have not considered any other factors in their lives or what would have happened without the project or how they might have been influenced by the group. Results may be affected by lots of other issues (e.g. they may have got married or won the lottery in that time!) that have nothing to do with the organisation so we cannot be certain the project is responsible for any changes we observe.

Happy Tracker

A baseline and review tool

- Happy Tracker

Happy Tracker is a very simple and easy tool developed by the Story Museum working with the New Economics Foundation (nef) to assess individual wellbeing. Participants simply use their body to indicate their level of wellbeing at points in time without the need for surveys or forms. Responses can also be verbal or written, whatever suits the situation.

When should we use it?

Happy Tracker can be great if you're working with a participant group who may be intimidated by more formal data collection methods or if you don't have the time or resources to use a more sophisticated approach.

The tool should be used before involvement in your museum or project to capture a baseline (or starting point) and then repeated either at the end of the involvement, or at regular intervals to be able to track changes in levels of wellbeing.

	How happy were you yesterday?		
	Before	After	Change
Participant 1	1	6	5
Participant 2	4	6	2
Participant 3	3	6	3
Participant 4	4	5	1
Participant 5	6	5	-1
Participant 6	7	7	0
Participant 7	3	6	3
Participant 8	4	6	2
Participant 9	2	5	3
Participant 10	3	5	2
Participant 11	4	5	1
Participant 12	4	7	3
Average change			2.0

How do we use it?

- Ask participants to stand in a line
- Ask participants to use their arms to indicate their answer to the question 'How happy were you yesterday?' on a 1-7 scale
  - For 1 – 'Not at all' - arms down by their sides.
  - For 3 – 'Neither happy nor sad' - arms straight out in front (parallel with floor)
  - For 7 – 'Completely' - arms straight up above their heads.
- Record the arm position of each participant (translated onto a 1-7 scale) and save in a spreadsheet against each participant
- Repeat the exercise at the end of the project (or at regular intervals, for example, monthly or quarterly) and record the result for each individual

- Minus the before score from the after score to reveal the movement on the scale for each participant. A positive number shows positive movement, a negative number shows negative movement

Add up the totals for all individuals before and for after to reveal the movement of the group overall.

You could ask more than one question if you wanted to understand more about people's feelings and emotions. For example, you could ask 'Overall, how satisfied are you with your life nowadays?' or 'How much do you care for your environment or surroundings?'. You could have a look at the LIFE survey for suggestions for other questions you could include.

Usable versus robust?

How useful and empowering?

The tool asks individuals to consider how they feel. This demonstrates you are interested in them and their wellbeing, which can empower individuals.

Happy Tracker requires little effort is very friendly and easy to use. You don't need any resources other than time and a pen and paper (although it can be improved with extra resources – see below). It is important to be aware that the tool is open to many biases that may affect how people report their wellbeing, for example, feeling they should report higher levels, worrying what other people will think, being influenced by others, or being overly affected by irrelevant circumstances, such as the weather. Using an arm position requires the person recording the response to judge which number on the scale is being represented that could result in misinterpretation.

The tool requires minimal organisation and record-keeping so it's useful if you don't have the capacity to use other measurement methods. Happy Tracker provides an idea of how the self-reported wellbeing of individuals is changing over time and does not require participants to do any form-filling or interviews, which may not be possible with some groups.

How can robustness be maximised?

- Make sure you collect data at the very start of the project so you have a baseline
- Make sure the wording of each statement is available on a slide or flipchart so people don't have to remember it
- Provide exactly the same guidance to participants each time and don't say anything that may lead their response
- Remind people they should be honest and open
- Provide diagrams or model the relative arm positions for the 1-7 scale as guidance
- Ask people to line up in alphabetical order (each time) and take a photograph to increase the accuracy of recording answers
- Ask people to close their eyes so they're not able to see each other's answers so that they may feel more comfortable to be honest.
- Always conduct the exercise at the same time of day and on the same day of the week

What can the results tell you and what do they not tell you?

The results can provide an indication of the change in wellbeing for individuals whilst they are engaged with your project. It is important to remember the possible biases in how they report as well as the caveat that we have not considered any other factors in their lives or what would have happened without the project. Results may be affected by lots of other issues (e.g. they may have got married or won the lottery in that time!) that have nothing to do with the museum so we cannot be certain the project is responsible for any changes we observe.

Review for personal or organisational change

Embedded Evaluation

A review approach

- Embedded evaluation

This is the idea of making evaluation part of the project itself rather than something that takes places outside of the project or at another time.

Here are some examples of embedded evaluation:

Beaney's Paper Apothecary project asked community groups to devise 'wellbeing prescriptions' using the museum displays, then dispensed these to the public from a two week installation apothecary built from recycled paper. The paper prescriptions asked visitors to record any side-effects, comments and reactions to their cultural treatments via a tear off slip to post into the feedback box.

The Story Museum gave people 'happy labels' to write comments on, and tie onto artefacts or the architecture. A similar technique has been used by youth clubs asking for feedback on their venue. Visitors are given red and green flags or post it notes and asked to place them wherever they choose around the space to feedback their good and bad views.

Another approach developed by the **Story Museum** was the 'mood tree'. The museum used colour coded leaves which could be fixed to model trees placed at the beginning and end of the exhibition. Red, orange and green leaves indicated bad, so-so, and good mood. At the end of the day the leaves on each tree could be counted to see if there was a change.

Other commissions like Shakespeare's Birthplace Trust, used the same approach, but with Lego bricks, sticky stars, colour coded smiley faces, or mood buckets, attempting to find imagery that was relevant to the project.

When should we use it?

Embedded evaluation can help you to gather insights into the experience of your project in the moment and can tap into the creativity of your visitors or participants. Embedding evaluation requires creative thinking and careful design to find out what will work for your projects and participants as it is unlikely to be 'one size fits all'. An analyst and artists or designer working together can develop these approaches very effectively.

How do we use it?

The potential ways you could embed evaluation techniques into your project are broad and will be specific to the design and structure of your project. A key principle of embedded evaluation is that it fits with your project and encourages participants to provide meaningful insights.

Usable and friendly vs robust?

How useful?

Embedded evaluation can be very useful to capture insights into the momentary experience or immediate reflection of your project. It may also be possible to theme and/or count the feedback afterwards for some quantitative analysis.

Embedded evaluation can be very empowering for participants especially if they see their contribution becoming part of the exhibition itself. Creative expression that inputs into project evaluation may be received in a more positive way than asking people to step outside of their experience and fill in a form.

Evaluation exercises of this nature can produce rich and valuable insights in a very friendly and creative way for participants. It is important to consider that results are subject to biases. People may feel they should say something positive or nothing at all (especially if feedback will be displayed publicly), opinions and thoughts may be influenced by what others think and have written, and people may focus on the enjoyment of the exercise itself and may not fully reflect on their experience at the museum.

How can robustness be maximised?

- Remind people that they should reflect on their whole experience in the exhibition or museum
- Remind people they should be honest and open (and that you welcome negative feedback as well as positive)
- Do not cherry-pick results, make sure you include all feedback
- Analyse the results using narrative evaluation to draw themed conclusions
- Combine with other research on a 'causal chain'. For example Manchester Museum know from academic research that play enhances wellbeing so their embedded approaches look for play (not wellbeing)

What can the results tell you and what do they not tell you?

The results can provide an indication of how the museum can make people feel, and provide some insightful qualitative data on their experience. Some quantitative data can be generated but should be treated as an indication of effect only. Results will not be scientific evidence of the impact of your museum for the reasons discussed above.

Narrative Evaluation

A review/analysis approach

- Narrative evaluation
- Narrative evaluation template

You may gather qualitative written feedback through forms, interviews, visitor comments, embedded evaluations and so on. There are ways to collate this data and include it as evidence in your evaluation and learning processes. Chiltern Open Air Museum have used this to evaluate their project, and other museums have used this approach to retrospectively analysis their visitor book and compare with results after a particular project.

When should we use it?

If you gather narrative data through various means and wish to find a way to collate it and use it to inform reflections on what you have achieved.

How do we use it?

There are different ways to analyse narrative data.

- **Word clouds**: A simple way to analyse a selection of text is to enter it into http://www.wordle.net/create or in software such as Nvivo. This will give you an impression of the how many times certain words appear in the text in the form of a word cloud.
- **Thematic assessment**: You can group sections of written material by assigning each to a theme in an Excel spreadsheet. You may wish to attach or 'code' sections of feedback with tags, for example, 'creative', 'new', 'positive' or 'negative'. You can then set up some simple formatting that reveals how many times a theme appears. There is a excel template on our website that will help you to do this.

Observation	Emotion	
	excited	2
	quiet	2
	fun	1
	negative	0
	enthusiastic	0
	creative	0
	imaginative	2
	VSA observation	1
One little girl picked up the smallest spider and carried it around before placing it carefully on the floor.	imaginative	
Another child shouted 'don't wake the spider!'	excited	
A group of children started jumping on the bubble wrap shouting 'is it real?'	imaginative	
One parent said to his children 'don't wake the spider' so they tiptoed towards it had a look and then carefully walked away.	quiet	
Several children walked into the area with parents and tip toe over the bubble wrap to see the display cases- it appeared they thought they shouldn't walk over it.	VSA observation	
With a little encouragement two children tried to walk around the spider without making any noise, competing with each other.	quiet	
Towards early afternoon many children taking pieces of bubble wrap to jump on- no longer trying not to wake the spider.	excited	
A number of children sitting amongst the bubble wrap and popping it.	fun	

Usable and friendly vs robust?

How useful and empowering?

These techniques can enable you to analyse and quantify big chunks of text or narrative quickly and easily and it can be useful to bring existing written data into the evaluation and to identify illustrative quotes.

How empowering the tool is depends on what is done with the results. If they are displayed publicly and people know they contributed to them they may find it empowering to see their words on the wall.

The approaches mentioned will support you to analyse narrative data in simple yet useful ways. Obviously how you frame and collect the data will have a considerable influence on the content of it, for example, if you ask people what they liked about a project the text will not offer a balanced view of the experience. Even if you do ask people to be honest, certain biases come into play which prevent them doing so.

How can robustness be maximised?

- Include data from as many sources and people as possible (and is relevant) and make sure to capture a range of views, for example, negative as well as positive feedback.
- Do not 'cherry-pick' results
- Use feedback that has been provided privately, for example comments cards posted in a box rather than comments written in a book

What can the results tell you and what do they not tell you?

The results can reveal prominent sentiments and feelings or pick up dominant themes from a section of text. Results won't tell you about anything that is missing or hasn't been mentioned. The results can provide a snapshot of analysis rather than an overall assessment of change but may back up other information, or tell you where to focus your evidence collection next.

Observational Evaluation

A review/analysis approach

- Observational evaluation

Observational evaluation involves researchers observing participants in projects to record their activities, behaviours, emotions, interactions, themes and so on, to get a sense of how people experience your project. It can also be an opportunity to engage the front of house teams.

When should we use it?

You should use it when you want an alternative and additional measure of experience and evidence that is not directly from the participant, or a particularly light touch approach. It can happen during a project activity and can be repeated to identify changes over time.

How do we use it?

The approach involves agreeing a list of behaviours, emotions or interactions you are looking out for (e.g. smile, creativity, sharing) and at agreed intervals recording which of this list can be observed. There are variations on this approach such as interval observation, frequency observation and blow by blow account. There is more about these different approaches in our detailed guidance. There are also formal observation approaches such as Dementia Care Mapping, which needs to be undertaken by trained observers.

Usable versus robust?

How useful and empowering?

Observational evaluation can be useful to provide another perspective of the experience of your participants that does not come directly from them. It can enable a focus on a specific area of interest, for example, if you are interested in how much intergenerational interaction is taking place or how much people laugh during an activity. Sometimes it might be the only evidence collection approach open to you.

Observational evaluation is not likely to be empowering for those being observed because to be truly robust participants shouldn't actually know it is happening, as we wouldn't want the observation to affect or change their behaviour. However, there is also an ethical issue which organisations need to resolve for themselves. Our guidance would include:

- ensuring only previously agreed behaviours are collected
- not making assumptions about demographic information (such as race, age or gender)
- including signage to the effect that observation is taking place.

It can, however, be empowering for front of house staff by involving them in the evaluation.

Observational evaluation requires organisation and effort to prepare but requires nothing of your participants. It involves subjectivity in both drawing up the list of items that you decide are important and the researcher's perspective as to what translates into a tick on the list. The narrow focus and subjectivity allow for many biases but being aware of these can capture some useful insights about your project. The approach doesn't tell you what your project is responsible for, for example, you may be observing someone who smiles all the time regardless of whether they are involved in your project or not.

How can robustness be maximised?

- Ask a range of people to input into and agree on the list of criteria and ensure you allow for positive and negative findings. You might for example, include a psychologist in this work
- Ask multiple researchers to observe the same situation to enables comparisons between findings in testing phase
- Ensure your observations are random (e.g. every five minutes) rather than driven by what you see and do not purposefully omit anything
- Observe a similar group of people (a control group) doing something else
- Combine with other research on a 'causal chain'. For example Manchester Museum know from academic research that play enhances wellbeing so their embedded approaches look for play (not wellbeing)

What can the results tell you and what do they not tell you?

Results can reveal elements of your participants' experience and can be objective if sufficiently validated through testing. We do not know if your project is behind the observed behaviours. Significant events could easily be missed due to narrow focus of the approach, for example, someone could experience a significant improvement in confidence but this may be difficult to categorise in a list of observations or may not be a particular focus of the project.

Valuation DIY

A review/analysis/valuation approach

- Valuation DIY

Valuation DIY is an empowering way for participants to discuss not just the extent, but the value of the different outcomes they have experienced through your project.

When should we use it?

Valuation DIY is to be done after an event to find out the relative value placed on the outcomes by participants.

How do we use it?

- Show the participants the outcomes of the project and seek agreement that these were achieved (at least to some extent)
- Put each outcome on a separate card and ask the group to rank the outcomes from most important to least important
- People are likely to have different views but encourage people to try to convince each other and reach a consensus
- Introduce cards with images of a selection of market goods i.e. things that people pay for, for example, a phone, a holiday, an adult education course, a bottle of wine.
- Ask people to slot these market value cards in between the outcomes as they have ranked them. This provides some guidance on how much people would be willing to pay for these non-market outcomes in a world where you would be able to buy them.

Usable versus robust?

How useful and empowering?

Valuation DIY is a useful way to find out which outcomes are most important to participants and to support them to attach proxy values to non-market outcomes.

Valuation DIY may empower people if they believe that they are being consulted, their voices are being heard and the results will inform the evaluation of the project.

This should be a fun and engaging exercise for participants to get involved in and shouldn't take much organisation or planning. The approach provides an indication of the relative importance of outcomes to participants but the valuation exercise should be approached carefully. There is an important difference between price and value (as well as cost). The price of the market good is not the same as how much people value it. Some people may value something very much and be willing to pay more than the market price, the market price represents the least they'd be willing to pay. For example, a fillet steak may cost £15. Someone who really enjoys a good quality steak may be willing to pay £20 for it, whereas a vegetarian wouldn't be willing to pay anything for it. So, the valuation exercise may encourage discussion but the results should be approached with caution.

How can robustness be maximised?

- There will be an optimum number of people to include to ensure a broad range of views without it being too many people to manage. The optimum number is probably around 5 or 6.
- The results can be checked against other research into valuations, for example those collected at the Global Value Exchange.

What can the results tell you and what do they not tell you?

The results can tell you the relative importance of the project outcomes to your participants which can be really useful when revisiting your project design and focus. The results do not tell you the magnitude of change experienced across the outcomes or whether what people say is the most important outcome aligns exactly with their experience. For example, an individual may report that new skills learnt at a workshop were the most important outcome for them whereas actually it was the friends they made at the workshop that had the biggest impact on their day-to-day life.

Which methods should I use? How do they compare?

Hopefully you are now aware of the strengths and weaknesses of the different tools in this toolkit. The following points should help to you decide which tools you should use in your project.

- Use the LIFE survey where possible as it will result in the most robust and meaningful measure of change. Your results will support a compelling case to funders/government
- Use less robust measurement techniques where this is not possible but be aware of biases and take measures to minimise the impact of them
- Use the qualitative or narrative methods to put the flesh on the bones of the changes you have identified and to tell the story and include the human interest side and case studies or quotes. This will help you to tell the story of your project (alongside the measured change) to Boards, participants and for publicity
- Make sure you use results to inform what you do to improve your projects. There is no point doing all the exercises, data collection and analysis if you don't then use the results to improve what you do
- Before you decide which approaches to use, consider the resources you have available, the purpose of your measurement (primarily for internal decision-making? reporting? external use? or to empower?), and the audience (is it your participants, your Board, your funders – or all three?)

This guidance was included in Happy Museum's contribution to 'Towards Plan A' for the Arts Council and RSA.

	Showing value and accountability	Learning and improving	Motivating and empowering
Funders and commissioners	Against targets Against others	Context of Social Value Act, social finance and payment by results	Context of austerity and natural limits, wellbeing and restructuring public services
Delivery organisations		Over time Against opportunity	
Practitioners and participants		Context of improvement and trade-off	For co-production

Tool table

Click here for the tool table.

Glossary

Baseline – data gathered before an intervention or project to use a starting point from which to measure change.

Controls for – controlling for a variable simply means taking out its effect through statistical methods.

Control group – a group of people who are not involved in your project but are similar to your participants from which you collect data to give you a point of comparison.

Counterfactual – what would have happened in the absence of your project or intervention 'what would have happened anyway'

Life satisfaction – a broader measure of wellbeing that picks up current mood and as well as an overall assessment of one's life is going in comparison to one's hopes and in comparison to other people.

Market value – something that is bought and sold and so has an agreed price or monetary value

Non-market value – something that is not bought and sold and so a monetary price or value does not exist

Review – is used here to mean the evidence collection that happens after a project or activity

Valuation – converting an outcome measure onto another metric that tends to be a monetary value (but can be on a non-monetary scale).