PBE Economics to improve lives

SPIRIT OF 2012 INVESTING IN HAPPINESS

# Helping funders measure what matters

## Using wellbeing to understand the value for money of Spirit of 2012 grants

By Sara MacLennan
and Jon Franklin

May 2025

PBE Economics to improve lives

We use economics to improve lives. Through analytical expertise and our close connection with the social sector, we help charities, funders, firms and policymakers tackle the causes and consequences of low wellbeing.

Our analysts, researchers and economists work on a wide range of issues related to low wellbeing, including mental health, education, employment, financial security, poverty, disability, inequality, volunteering and civil society. Working with over 600 volunteer economists, we have supported over 600 charities since 2009.

SPIRIT OF 2012
INVESTING IN HAPPINESS

Spirit of 2012 is the London 2012 social legacy organisation.

It was founded with a £47million endowment from the National Lottery Community Fund in 2013 and will undertake a planned closure in 2026.
Their research, learning and insights from over a decade of funding projects to inspire a social legacy is available at www.spiritof2012.org.uk. learners.

# Summary

**Funders influence how impact is measured in the social sector. Their data needs shape both the type of data charities collect and how they use it. However, the need for charities to balance data collection with providing services to beneficiaries can create tension between funders and grantees. Funders must find a balance between asking for data to understand impact and reducing the burden on grantees and their beneficiaries.**

Wellbeing measurement offers a valuable opportunity for funders wanting to become more measurement focused. Our wellbeing – the sense of how we are feeling and functioning – is now a robust outcome that can be measured using simple, pragmatic tools. These measures can help to provide robust impact conclusions while minimising costs and complexities for funders and their beneficiaries.

Spirit of 2012 – the legacy funder from the London 2012 Olympic and Paralympic Games – adopted the Office for National Statistics' (ONS) standardised personal wellbeing questions across its grantees in order to measure what mattered most to the stakeholders it worked with. This helped to build the evidence base for how sporting, cultural, and community events across the UK can improve wellbeing, inclusion, and social connection.

Spirit of 2012 used the data to great effect. It worked with grantees to understand who they were supporting and refine programmes throughout the delivery period to ensure they were targeting those that needed the support the most. Spirit of 2012 used the data to identify comparison groups for some initiatives to better understand what would have happened in the absence of the interventions it was funding.

As they approached their planned closure, Spirit of 2012 wanted to explore what more could be learnt from the data and reflections that they had collated from grantees over the past 12 years. PBE (Pro Bono Economics) was commissioned as one of three Legacy Learning Partners to review archival content and use it to produce new insights for their own audiences. Having gathered data using the ONS' standardised personal wellbeing questions, it opened up the opportunity of doing retrospective wellbeing cost effectiveness analysis on a selection of closed projects. Three former grantees were approached, selected to provide some sense of the variety of what Spirit of 2012 has funded, with all three agreeing to take part.

Through this work, Spirit of 2012 has gained value for money assessments to help demonstrate the difference that fun and supportive 'free time' activities make to our lives. This was particularly important for Spirit of 2012 as these participatory sport, arts, and volunteering projects typify those that are used in planning for and delivering a legacy from events.

## The key findings include:

### The EmpowHER programme

an initiative led by UK Youth, in partnership with the British Red Cross and Young Women's Trust, which engaged young women and girls in inclusive, meaningful social action – represents good value for money, delivering around £5 in wellbeing benefits per £1 spent.

### Get Out Get Active (GOGA)

a programme led by Activity Alliance that supported disabled and non-disabled people to enjoy being active together – delivered around £3.70 of wellbeing benefits for every £1 spent.

### City to Sea

was a surf-therapy project which targeted young Londoners whose complex physical and mental health needs made them vulnerable to isolation. Spirit of 2012 funded Laureus Sport for Good to work with The Wave Project to deliver this initiative. PBE analysis suggests less than a third of the improvement in wellbeing experienced by participants needs to be down to the programme in order for the benefits of the programme to outweigh its costs.

We can also learn a lot from Spirit of 2012's experiences. First, there are barriers that can make it difficult for grantees to measure wellbeing – and a funder can help to bridge the gap. Secondly, developing an understanding of what might have happened in the absence of the programme can be particularly important where programmes are having a 'protective' impact (preventing things from getting worse). Finally, providing central support for grantees to access the right expertise can be critical to deliver a truly 'measurement focused' approach. Our report captures these important lessons and offers guidance via three of Spirit of 2012's grantee experiences or 'case studies'.

Ultimately, Spirit of 2012's experience demonstrates that putting measures of wellbeing at the centre of impact measurement approaches can deliver pragmatic, insightful results that could form a useful blueprint for other funders to follow.

# Funders face a difficult trade-off when it comes to impact measurement

**Funders play an important role in shaping impact measurement practices in the social sector. The choice of what data to request from grantees throughout the lifecycle of a grant, as well as the requirements for evidence of broader social impact at the application stage, can have a significant influence on what data charities collect and what they do with it. Previous reviews have highlighted that "perceived pressure from funders" is the most significant motivation for organisations to decide to start measuring their impact.[1]**

This desire for more and better impact data appears to be getting stronger over time.[2] Nine out of every 10 funders believed that impact measurement data can make both charities and funders more effective.[3]

Although it may not be a definitive factor on whether to support charities in the first place, research suggests it does impact on grant renewal and can influence the review and design of new grant-making programmes. There is also growing evidence of funding organisations wanting to understand and quantify their own impact, likely to secure long-term financial support for their own organisation or influence wider systemic change.[4]

However, these demands for more impact data can create tensions between funders and their grantees. Balancing the need to gather data with the practical demands of delivering vital services to beneficiaries can result in data only being collected 'symbolically' but never really being used to inform decisions. It can also lead to duplication of efforts as different data is needed for different funders, alongside internal needs.[5]

These tensions have led to some push-back against funder-imposed measurement frameworks, with a stronger focus on 'trust-based' approaches that place minimal demands on grantees. This empowers grantees with maximum flexibility to use their expertise on how to best solve the problems they are tackling and measure the impact of that in their own way.[6]

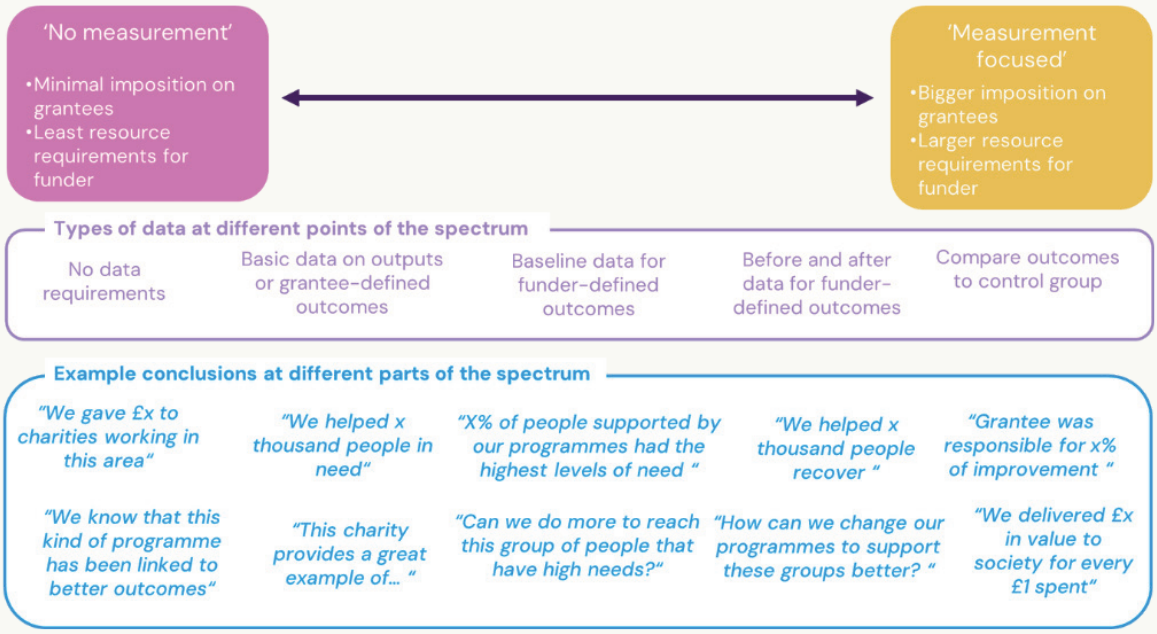
(1) Harlock (2013): Impact measurement practice in the UK third sector: a review of emerging evidence. Accessed March 20, 2025
(2) PBE (2023): Unleashing the power of civil society. Accessed March 20, 2025
(3) Kail et al. (2013): Funding Impact – Impact measurement practices among funders in the UK. Accessed March 20, 2025
(4) Chaidali P et al. (2024): Charity impact reporting: Informing the forthcoming Charities statement of recommended practice. Accessed March 20, 2025
(5) Chaidali P et al. (2024)
(6) Salehi (2024): The rise of trust-based philanthropy. Accessed March 20, 2025

This means that funders face a difficult trade-off between placing measurement demands on their grantees, to more clearly and consistently understand their impact, and minimising the 'burden' placed on grantees and their beneficiaries. We like to picture this trade-off on a spectrum – moving from 'no measurement' on the left-hand side, through a range of 'light touch' approaches up to more intensive 'measurement focused' approaches on the right-hand side – as shown in Figure 1.

If a funder chooses to position themselves on the left-hand side of this spectrum they impose minimal requirements on their grantees and need the least resources internally to use this information. However, the strength of conclusions they can draw about the impact of their grantees will be the weakest – likely to be limited to knowing what sorts of causes were supported and pre-existing evidence on how that helps to achieve end goals. If a funder chooses to position itself on the right-hand side of this spectrum it will result in higher costs. These costs may fall to themselves as a funder, if they cover the costs in full, or be imposed on its grantees if there are measurement expectations without specific funding. The funder will need more skills and expertise to make the most of the evidence generated. On the plus side, it can generate far stronger conclusions from this evidence, helping funders to understand what works for tackling the challenges they're interested in and the value for money delivered by their investments.

There are, of course, middle ground 'lighter-touch' options that can provide valuable insights for funders and their grantees while avoiding some of the largest costs. For example, agreeing a pre-defined measure of 'need' to ask at the beginning of an intervention can help to evidence that the programmes are reaching those with the greatest needs. This can then be used to support a conversation about whether there are any key groups that are being missed. If this can be extended to gathering follow-up data after a grantee has provided support, then it can be further strengthened by checking that similarly positive outcomes are achieved by different groups of end beneficiaries. This can prompt further reflection and qualitative insights that can help to evolve and improve project delivery.

Figure 1: Funders face a difficult trade-off when it comes to impact measurement




There is no 'right' place for funders to sit on this spectrum, as long as their expectations of what they want from the evidence aligns with the approach they choose to implement. But this needs to be a conscious choice from the beginning. The 'right' questions need to be asked – often starting before the programme or project has started. Failing to plan for this can limit the options for building evidence at a later date. For example, value for money analysis can rarely be done as an 'add-on' at the end of the project if it has not been planned in from the very beginning of monitoring and evaluation.

This choice of where a funder should sit on the spectrum will be impacted by a number of considerations, such as the evidence needs of the funder, the strength of evidence that already exists, the size of the grants, the capabilities of the grantees, and the 'organisational philosophy' of the funder. As recent guidance from New Philanthropy Capital highlighted, the key is to be intentional and proportionate with the approach.(7)

However, given this choice, a key question for many funders is whether there is an approach to measurement that can provide as many of the benefits from a more 'measurement focused' approach while minimising the costs for grantees and funders. While there is no silver bullet, we have found that wellbeing measurement provides an excellent opportunity for funders that want to become more measurement focused. We explore this further throughout this report.

(7) New Philanthropy Capital (2024): Monitoring, evaluation, and learning with trust and equity: a guide for grant-makers and philanthropists. Accessed March 20, 2025

# Wellbeing offers a valuable opportunity for funders wanting to become more measurement focused

Our wellbeing – the sense of how we are feeling and functioning – is now a robust outcome that can be measured using simple, pragmatic tools. The science of wellbeing measurement has evolved rapidly in the UK over the last two decades. A key milestone was when the ONS started measuring the nation's personal wellbeing using a standardised set of questions in 2011. These questions have been rolled out to a wide range of government-funded surveys and supported the rapid development of an ever-expanding literature on what influences the overall quality of people's lives and what should matter for policy.[8]

The credibility of these new approaches has been underlined through its adoption in the government's official guidance for developing policy.[9] This guidance, issued by His Majesty's Treasury (HMT) in 2021, highlights the role that government policy has in increasing the wellbeing of the population and emphasises the role that wellbeing measurement can play in understanding the impact of policies. It has helped to establish wellbeing as a reliable and convincing tool for understanding the impact of social programmes.

## ONS4 personal wellbeing questions

The ONS uses four survey questions to measure personal wellbeing.
These are known as the ONS4.[10]

### The questions are:

- **"Overall, how satisfied are you with your life nowadays?"**
- **"Overall, to what extent do you feel the things you do in your life are worthwhile?"**
- **"Overall, how happy did you feel yesterday?"**
- **"Overall, how anxious did you feel yesterday?"**

(8) See for example Layard and De Neve (2023): Wellbeing: Science and Policy as well as yearly World Happiness Reports

(9) HMT (2021) Green Book Supplementary Guidance: wellbeing. Accessed March 13, 2025

(10) The ONS first added these questions to the Annual Population Survey (APS) in April 2011. The APS is the source of the national estimates of personal wellbeing in the UK that are published quarterly by the ONS.

People are asked to respond to the questions on a scale from 0 to 10 where 0 is "not at all" and 10 is "completely".

These questions ask a person to give their own, personal view of how they feel about their life: they are subjective measures. Subjective wellbeing measurement provides a direct and meaningful indicator of the quality of life of an individual, based on an individual's own assessment of what is most important and how they are doing. The measures have been widely used and tested for over 50 years[11] and are well accepted in academia, practice, and policy.

## So, why is wellbeing measurement helpful for funders?

Our experience working with charities and funders has highlighted a number of reasons:

- It's a simple and pragmatic way of capturing what matters to many social sector organisations and funders: fundamentally, the majority of charities exist to try and improve the quality of their beneficiaries' lives. Wellbeing provides an excellent indicator of the overall quality of life, and it does this using four simple questions that have been proven to be reliable and robust.

- It can be used in a wide variety of contexts: all too often funder measurement frameworks become unworkably complex, with an array of different outcome measures covering multiple different dimensions. This often confuses grantees and results in a complex array of data to unpick for the funders themselves. Using an umbrella measure of quality of life allows a single framework to be applied across lots of different dimensions.

- It provides a great way of demonstrating that charities are reaching those with the greatest needs: because national wellbeing levels are measured regularly by the ONS, it is straightforward to compare the wellbeing levels of charity beneficiaries to the wider public. This can demonstrate that funding programmes are reaching those with the lowest levels of wellbeing (and therefore the greatest need for support).

(11) And meet the four psychometrical properties of reliability, validity, sensitivity, and discriminant validity.

- It can be easier to get stronger conclusions about impact using comparison groups: because the ONS measures of wellbeing have been rolled out to so many surveys, it can make it far more feasible to understand what might have happened to the wellbeing of beneficiaries in the absence of the grants. Typically, this is done by identifying individuals from these datasets with similar characteristics (particularly initial wellbeing scores) and comparing the changes in wellbeing they experienced over time against beneficiaries for grantees. While this sort of evidence may not be as strong as large-scale experiments, it can provide a pragmatic way of reaching a relatively robust evaluation for many more organisations.

- It is an accepted way of measuring interventions: subjective wellbeing measurement is now an accepted measure, a credible way of measuring interventions, and increasingly used by central government.

- It allows charities to assess the value for money of grants: for many small charities or community groups, there are no financial outcomes that can be used to directly capture the monetary benefit of their activities and the social outcomes they generate. This often means that the evidence they generate following delivery cannot easily assess whether the monetised benefits outweigh the costs of the investments, to inform 'value for money' decision-making. Using the HMT's Cost Benefit Analysis guidance, the broad social impacts of an intervention can be captured and monetised using subjective wellbeing measurement (life satisfaction).[12] Organisations that are working to improve the real, lived experience of those struggling can make the case of the value that they bring, even where they may not lead to any changes in markets, changes to productivity, or savings in benefits.

(12) HMT (2021)

# Spirit of 2012's approach to using wellbeing to understand the impact of grantees

**Spirit of 2012 is the legacy funder from the London 2012 Olympic and Paralympic Games. It builds sustainable social legacies from the inspiration of events, investing to improve how people feel about themselves, other people, and their communities. Given this aim, it was of central importance to Spirit of 2012 to measure whether and how its investments improved how people feel.**

Spirit of 2012 recommended that all projects measured and evaluated the wellbeing of participants using the ONS' measures of personal wellbeing. This followed a lengthy process of engagement with 30 cross-sector stakeholders. The process identified a theory of change, with wellbeing as one of the core outcomes that all funded projects should seek to impact. Spirit of 2012 was committed to meaningfully measure wellbeing for participants in its programme. It wanted to measure what mattered to people, to measure how people felt about themselves, as an end in itself. It considered a number of different subjective wellbeing measures. The ONS4 were a natural fit for recommended questions, based on their simplicity, rigour, and widespread use.

> "It matters what you measure. We should measure what matters to people, assessed by themselves. The validated and consistent personal wellbeing questions from the ONS provide a straightforward, simple way of doing this across a wide range of different programmes"
>
> – Amy Finch, Director of Policy & Impact, Spirit of 2012

## Spirit of 2012 impact statement

**We believe that enabling people to participate in a wide range of inclusive activities, and engaging together in their communities, will:**

- Improve the wellbeing of individuals, communities, and society as a whole.
- Improve perceptions (including self-perceptions) and attitudes towards disability and impairment.
- Lead to greater social cohesion and understanding.

The aim of Spirit of 2012's approach was to be more measurement focused; enabling evidence of reaching those with the greatest needs, understanding change in outcomes over time, strengthening impact measurement using comparison groups, as well as assessing the value for money of some of the programmes supported. They wanted to test their theory of change on how events can be harnessed to leave a longer-term legacy and make the case for more intention, and sustained funding for events legacy programming. Ultimately, this was to build the case that investment in fun and supportive 'free time' activities, those most likely to be used in event legacy planning, makes a valuable difference to people's lives.

This section explores the key learnings and describes how the approach related to one of the three case studies, EmpowHER, in order to bring it to life. Spirit of 2012's work with EmpowHER, a programme led by UK Youth, in partnership with the British Red Cross and Young Women's Trust, provides a powerful example of how this approach to wellbeing measurement supported both organisations.[13] EmpowHER supported young women and girls to engage in inclusive, meaningful, youth-led social action in their communities.

(13) Further details of the evaluation work are provided in Annex A.

## Reaching those with the greatest needs

Spirit of 2012 placed a priority on improving wellbeing, especially for those with the lowest levels of wellbeing. Initially, the measurement approach helped Spirit of 2012 to work with grantees to explore baseline data measures of wellbeing at the start of the programmes. This included creating funding programme 'pauses' after an initial delivery run or 'cohort' of a programme to enable time to stop, assess, and learn from what its data and wellbeing analysis revealed. This helped them to adapt programmes and ensure they were reaching those with the greatest level of need.

This approach helped the EmpowHER programme to reflect on progress during the first cohort of the programme and pivot to ensure it reached those with the greatest need. UK Youth identified that a relatively small proportion of disabled young people were being reached in the first cohort of its work. This is a group that disproportionately experiences low wellbeing – the average disabled young person started the programme with a wellbeing score of 5.8 out of 10, compared to 6.3 out of 10 for those without a disability. UK Youth committed to recruit a greater proportion of disabled young people at the end of cohort 2. They achieved this through working with Autism and ADHD, a delivery partner specifically engaging disabled young people, as well as promoting an access fund to reduce the barriers associated with certain groups participating. Subsequent cohorts had a higher number and proportion of disabled people.

## Understanding change in outcomes over time

Spirit of 2012 worked with grantees to encourage and support follow-up with beneficiaries to see how outcomes changed over time. They encouraged discussions on where follow-up is relevant. For example, is there a level of engagement for which we would expect a change in life satisfaction? Typically, this would be where programmes have an ongoing relationship with someone over time where they might see a 'change in state', such as a change in feelings of loneliness, or a change in the immediate environment, or relationships at work. It is less relevant for a one-off event, where momentary or experiential measures may better capture changes.

This can require a clear explanation of what sorts of support and activities 'count' when programmes are delivered in different ways in different parts of the country. Where possible, being able to map participant survey responses to the location or activity type and whether they remained engaged can be useful.

This data was available for EmpowHER. Young women and girls completing EmpowHER experienced an average improvement of 0.95 life satisfaction points over a period of around 4.5 months. This is a large increase – to put it into context, the average increase in wellbeing for an unemployed adult finding a job is around 0.5 life satisfaction points. However, at this stage, not all of this change in wellbeing can be attributed to EmpowHER – some of it could have been down to other factors in these young people's lives.

## Enabling stronger conclusions about impact using comparison groups

**A key part of any impact evaluation is understanding what might have happened to participants in the absence of the intervention. This helps to check that changes seen in programme participants are additional – above and beyond what would have happened anyway – and can truly be attributed to an intervention. Identifying a comparison group – a set of very similar people that didn't receive the intervention – and comparing their outcomes against those involved in a programme is often the most pragmatic solution for many social sector interventions. As part of this process, Spirit of 2012 commissioned PBE to work with UK Youth to identify a comparison group from publicly available research surveys.**

The analysis highlighted that the EmpowHER programme was having an even bigger impact than first expected. In the absence of the programme, similar young women and girls typically experienced a decline in wellbeing of around 0.4 life satisfaction points. In other words, the programme was having a protective impact in addition to the boost in wellbeing seen among participants. This means that the total wellbeing impact that can be attributed to the EmpowHER programme is likely to be around 1.3 life satisfaction points.

# Assessing value for money of grants

**Value for money analysis provides an assessment of the value of benefits provided, compared to the cost of a programme. It can help to support grantees attract funding from other sources as well as prioritise between different types of programme.**

Supporting grantees to understand their value for money typically requires us to place a monetary value on benefits delivered and compare this to the costs. HMT guidance has established an approach to valuing changes in the ONS life satisfaction measure. However, there can be a number of complications related to understanding costs. For example: Is it possible to isolate sunk costs that are required for programme set-up but then fall for future cohorts? Are there finance costs that do not contribute to achieving outcomes for a particular cohort of individuals? Have central or overhead costs been included?

For EmpowHER, the analysis suggested the programme was good value for money – delivering around £5 in wellbeing benefits per £1 spent. PBE used the HMT valuations for wellbeing to estimate, in the central case, that the programme delivered around £6,500 of wellbeing benefits per participant. When compared to a cost of around £1,300 per participant, it can be seen that the benefits clearly outweigh the costs. Importantly, these findings remain robust to a range of different assumptions about the value of wellbeing and the persistence of the impact over time.

These powerful conclusions for EmpowHER demonstrate the value that a wellbeing approach can bring to measuring impact for funders and their grantees. However, as ever, it is not always as straightforward as this example suggests – there were important lessons from Spirit of 2012's experience that are explored in the following section.

# Lessons from Spirit of 2012's grantees' experience of wellbeing measurement

All organisations learn through doing, and over the course of development, funding, and evaluation, Spirit of 2012 has learnt which aspects of wellbeing measurement and analysis could have benefited from even further input or support, to enable the data to be gathered, used, and useful.

**In this section we review four key lessons for funders that emerged from Spirit of 2012's work:**

1. There are barriers to measuring wellbeing – and a funder can help to bridge the gap.
2. The benefits of identifying matched comparison groups.
3. Supporting access to the right expertise to support charities with impact measurement.
4. A future need to capture and incorporate wellbeing data from people with learning disabilities and neurodivergence.

# Lesson 1: The barriers to measuring wellbeing – and how funders can help

**Getting buy-in and agreement from grant holders for using any measurement framework is vital, and wellbeing measures are no exception.**

Spirit of 2012 started by setting an upfront expectation in application forms that wellbeing would be evaluated consistently with the ONS4, as well as including this in the terms and conditions of the grant. They took time to provide a clear explanation of why Spirit of 2012 wanted to know this information and how they were going to use it, as well as producing clear monitoring and evaluation guidance.

Spirit of 2012's grants included up to 10% allocation specifically for monitoring and evaluation activity, which could be allocated in-house, contracted, or spent in combination. Grant monitoring forms included a data collection template for wellbeing data – this was revised several times in response to feedback to ensure it was as user-friendly as possible.

Spirit of 2012 actively sought to build capabilities within their grantees through learning events and a learning support partner.

## Example of funder support for measuring subjective wellbeing

A mix of learning events and expertise to support grantees with evaluation

Spirit of 2012 provided a range of evaluation support. It varied by grantee but included:

- A funded workshop to create and develop their theory of change.
- Advice on what we know about improving wellbeing from the What Works Centre for Wellbeing, which could be built into design.
- Peer learning on wellbeing measurement and using the ONS4.
- Support and advice on measuring wellbeing. For example, the evidence that ordinary people thought talking about their lives and feelings was important; practical ways of managing distress.
- Funding for independent and expert evaluation (monitoring and evaluation could be up to 10% of the grant allocation, which could be in-house or externally commissioned).

It was critical to listen to concerns about the questions and provide reassurance.[14] A particular challenge was the perception that 'vulnerable people' shouldn't be asked subjective questions. However, highlighting the importance of having the voice of people with the lowest wellbeing and the highest barriers considered in the evaluation, as well as practical support from organisations who were already using the ONS4, helped to find practical solutions to these concerns. In short, understanding the position of grantees, listening to their concerns, and responding by providing appropriate training and support made an enormous difference to grantees' willingness to embrace a unified wellbeing measurement approach.

## Example of funder support for measuring subjective wellbeing

### Life satisfaction relevance

**It was important to discuss whether and when the life satisfaction question would be relevant. For example, some projects asked the life satisfaction question at the beginning and end of a day's event, for their 'pre' and 'post' measures. Although this may show an improvement in wellbeing over the course of this day, which is important in its own right, we cannot infer that there are lasting changes in overall feelings about their life, measured through life satisfaction.[15]**

Many programmes were designed specifically to change people's levels of activity, or feelings of belonging, through a number of week's engagement. It could be expected that such programmes would have a longer lasting change in wellbeing, for at least the length of engagement and in many cases, even further. A three-month or six-month follow-up, as is demonstrated in some of the case studies here, can test to what extent these benefits were maintained.

(14) For example, that the questions had been tested, that participants always have the right not to complete a survey/interview (or a particular question within it).

(15) WELLBYs are changes in life satisfaction per person per year, so a change for one full day would be the change * 1/365.

# Lesson 2: The benefits of identifying matched comparison groups

**Changes in wellbeing need to be understood in the wider context. Covid took place in the middle of many of the Spirit of 2012 investments, when subjective wellbeing dropped, especially around specific lockdown events. In these cases, 'maintaining' a starting level of wellbeing for participants could well be a positive impact.**

This is where matched comparison groups can play an important role in understanding the impact story of grantees. Essentially, an anonymised dataset with before and after wellbeing measures, as well as demographic characteristics, taken from programme grantees is used to identify a comparison group of similar people from publicly available survey data. The difference in wellbeing outcomes for those in the programme can be compared against those from the comparison group. This helps us to understand the additional impact achieved by the programme, above and beyond what might have been expected for similar people that did not receive support. This can be particularly important for 'protective' interventions that avoid wellbeing outcomes getting worse or during periods when the average level of wellbeing is declining substantially. It helps us to understand what might have happened had these beneficiaries not received the support they benefited from.

Although Spirit of 2012 encouraged matched sampling, most organisations did not do this. The difference in value for money conclusions possible where a matched comparison group is used, compared to conclusions without a matched comparison group, is best demonstrated through two case studies: Get Out Get Active, where a comparison group was identified, and The Wave Project, where it was not possible to identify a comparison group.

Get Out Get Active (GOGA), created and led by Activity Alliance,[16] was a programme that supports disabled and non-disabled people to enjoy being active together. GOGA aimed to get some of the UK's least active people moving more through fun and inclusive activities. They tracked grantees' wellbeing at different points in time and retained that data on an individual-by-individual basis. This allowed us to contrast their outcomes against a comparison group of similar people identified from a national research survey. Our weighted analysis demonstrated that disabled participants experienced a small positive increase in wellbeing of around 0.05 life satisfaction points (on a scale of 0-10). On its own this was not a significant change, but our comparison group typically experienced a decline in wellbeing of 0.4 life satisfaction points. In

(16) Funded by Spirit of 2012, with additional investment from Sport England and London Marathon Foundation.

other words, the programme was having a protective impact for disabled participants. When this impact is monetised, it suggests that GOGA was delivering around £3.70 of wellbeing benefits for every £1 spent.

City to Sea was a surf-therapy project that targeted young Londoners whose complex physical and mental health needs made them vulnerable to isolation. The social prescribing project was delivered by Laureus Sport for Good and The Wave Project over a three-year period between 2018 and 2021. While the project gathered wellbeing data for participants at the beginning and end of the project, by the time of this value for money analysis, the individual level data was not available to be able to track an individual's progress over time.[17]

This means that while we can see that average life satisfaction of participants was 0.8 points higher at the end of the programme compared to the start, it is not possible to find a matched comparison group. This limits the strength of what can be concluded about City to Sea's impact to, for example: break-even analysis – it is likely that less than a third of the improvement in wellbeing experienced by participants needs to be down to the programme in order for the benefits of the programme to outweigh the cost; or scenario analysis – if these benefits lasted for three months and the programme was responsible for 60% of the benefits then the programme would deliver around £2 of benefits for every £1 spent.

The difference between the headline conclusions for GOGA compared to City to Sea demonstrate the power of having a comparison group. While for GOGA the conclusions can be shorter, sharper, and easier to communicate, for City to Sea they need to be more carefully caveated and do not have the same immediate impression on readers.

Figure 2: The difference a comparison group can make to the strength of conclusions

Project with a comparison group:

"GOGA delivered around £3.70 of wellbeing benefits for every £1 spent"

Project without a comparison group:

"Less than a third of the improvement in wellbeing experienced by City to Sea participants needs to be down to the programme in order for the benefits of the programme to outweigh the cost spent"

(17) The matched data was available at the time of the project, but due to data anonymity regulations, the specific dates and identifying characteristics from the completed questionnaires had to be deleted after five years. At the time of the project, only the average figures were shared with Spirit of 2012.

# Lesson 3: Supporting access to the right expertise to support charities with impact measurement

While one of the benefits of using wellbeing measures is their simplicity and accessibility, there are some elements of impact measurement that require specific technical expertise.

## For example:

- Providing support for quality assuring data collection: this can include initial advice on the approach to gathering data, as well as reviewing data after it has been collected. Common challenges relate to issues such as assessing the best approaches to sampling, and testing that data collected is representative of broader cohorts of beneficiaries being supported.
- Supporting with analysis: typically, it takes specialist skills and knowledge to identify comparison groups or provide a value for money assessment. This can be a vital capacity gap for grantees and it is unrealistic to expect them to deliver this in-house.
- Provide support with interpretation of evidence: understanding how to use and interpret subjective wellbeing data can be challenging for organisations. For example, helping them to understand what change in wellbeing is 'good' and how it compares to other, similar interventions.

Spirit of 2012 highlighted the need to commission and deliver this support centrally to deliver the most 'measurement focused' approaches to impact monitoring.

# Lesson 4: A future need to capture and incorporate wellbeing data from people with learning disabilities and neurodivergence

**The ONS4 questions are not accessible for some people, including those with severe learning disabilities and neurodivergence. At the same time, these people may be some of the target group(s) of a particular intervention, for whom it is of particular importance to hear their views and understand the potential impact.**

Spirit of 2012 encouraged discussions and shared learning of how wellbeing could be measured in an accessible way. For example, one grant holder developed an 'easy read' version of wellbeing questions and also spent time using the questions qualitatively with their beneficiaries to understand what made them happy, what made things they do worthwhile. Others developed attractive visual resources to aid comprehension.

However, data from these amended versions is not yet possible to translate to monetised wellbeing for use in cost benefit analysis, since there is not yet a translation between life satisfaction on a 0-10 scale[18] and alternative, simplified versions. Translations are available between other wellbeing measures,[19] but not visual scales[20] or others which may be suitable for people with learning disabilities. This is an important area where further support and guidance, for example from academia, would be welcome, both on asking questions of people with learning disabilities and neurodivergent people, but also on the translation to life satisfaction (WELLBYs).[21] This will enable the voices of all groups to be incorporated in wellbeing value for money analysis.

(18) A change in life satisfaction is the specific measure which can be monetised for use in cost benefit analysis according to the HMT (2021) guidance.

(19) See Parkes (2025) Mapping Functions for Wellbeing Measures to Generate WELLBYs for Use in Economic Evaluation. Centre for Economic Performance Discussion Paper (Forthcoming).

(20) For example, scales which use smiley or sad faces to express the meaning of the different levels, rather than wording and numbers.

(21) A one point change in life satisfaction measured on a 0-10 scale, for one person for one year, is referred to as a WELLBY.

# Conclusion

**Funders face difficult trade-offs when it comes to setting measurement frameworks for their grantees. There is no 'right' answer, but decisions on the data to gather from grantees need to be clear, purposeful, and proportionate.**

**Wellbeing measurement offers a powerful opportunity for funders that want consistent outcome measures across a range of different interventions and a quicker, easier way to estimate value for money.**

However, to make it work in practice requires leadership and support from the funder. They will need to encourage the use of standardised wellbeing measures and work in partnership with grantees to enable all different types of organisations to use the measures. Peer-to-peer learning can play an important role in this process, as well as designing in moments to 'pause' delivery to reflect on data and use it to shape programme delivery. It is important for funders to consider the need to commission specialist support to complement the skills that they and grantees already have.

Spirit of 2012's experience, and the case studies outlined in this report, demonstrates that putting measures of wellbeing at the centre of impact measurement approaches can deliver practical results. The wellbeing value for money analysis that was possible helps to demonstrate the value of their investments in fun and supportive 'free time' activities: the difference which these investments made to people's lives, as assessed by themselves. And importantly, that this was 'worth it': in each of the case studies, we estimate that the monetised value of the wellbeing impact is likely to outweigh the financial costs. Each of these case studies may not have increased traditional measures of economic output, but they made a real difference to people's lives.

Citation: If you are using this document in your own writing, our preferred citation is: PBE. 2025. Review of Helping Funders to Measure What Matters. Pbe.co.uk. 15 May, 2025. www.pbe.co.uk /research-policy/insights/helping-funders-measure-what-matters.


# Annex A: The Wave Project – City to Sea

## Summary

City to Sea was a surf-therapy project which targeted young Londoners whose complex physical and mental health needs made them vulnerable to isolation. Participants were referred to City to Sea by Children and Adolescent Mental Health Services (CAMHS). The social prescribing project received £0.4 million of funding from Spirit of 2012 and match funders. It was delivered by Laureus Sport for Good and The Wave Project over a three-year period between 2018 and 2021.

At the time, and as a small charity, The Wave Project was at the early stages of embedding wellbeing evaluation across all their programmes.[22] As such, there is insufficient data to complete a robust economic cost-benefit analysis. It is a useful case study of the conclusions which can be drawn when more limited information is available – even though there is no robust 'benefit cost ratio', we can carry out break-even analysis and communicate value for money scenarios.

(22) In the following five years, The Wave Project further developed its evaluation process, as described in the technical annexes.

## We found that:

- With an average participant wellbeing of 6.3 at the start of the project, compared to a national average of around 7.6 for young people aged 10-17,[23] The Wave Project was effectively targeting those with lower wellbeing.[24]
- On average, wellbeing was reported 0.8 life satisfaction points[25] higher after the six-week Surf Therapy course. A strong caveat is that the reported wellbeing before and after were not necessarily from the same individuals.[26] We could have a bias if those seeing the biggest improvements were more likely to have responded to the follow-up survey.[27] In addition, we cannot be confident how much can be attributed to the programme as opposed to other changes occurring in people's lives. The duration of benefits is also uncertain. Wellbeing was only measured at the end of the six-week programme, but the benefits could have persisted for over three months, based on evaluations of other similar programmes.[28]
- With the information available, we carried out a wellbeing 'break-even analysis', where we explore these uncertainties in scenarios.
- Despite the limitations in the evidence available at the time and resulting uncertainties, we estimate that City to Sea is likely to offer good value for money across a range of scenarios.

(23) Not exactly the same age group. The Wave Project targeted those aged 8-21, compared to a national average of those aged 10-17.

(24) National average for young people aged 10-17 was 7.6 across the timeframe (The Children's Society annual Household Survey).

(25) On a 0-10 scale.

(26) Due to data anonymity regulations, the specific dates and identifying characteristics from the completed questionnaires had to be deleted after five years, making it not possible to trace this through.

(27) The view from The Wave Project was that the response to follow-up was often related to parental involvement, rather than level of wellbeing change.

(28) Devine-Wright, H. & Godfrey, C. (2020): The Wave Project: Evidencing Surf Therapy for Young People in the U.K. Accessed Feb 26, 2025Other studies of group therapy which combine surf therapy with cognitive behavioural therapy show benefits maintained at three months.

This case study demonstrates that wellbeing cost effectiveness analysis can still be a useful tool, even for projects and programmes that find themselves in the common position of not being able to gather before and after data from the same participants, or create matched comparison groups. The Wave Project has since started an evaluation programme which will monitor wellbeing with a significantly longer follow-up as well as matched before and after data. However, even if organisations are only able to measure life satisfaction before the programme, then this data is helpful. For example, City to Sea could confirm that they were effectively targeting those with lower wellbeing.[29]

### Measuring wellbeing lesson 1: understanding beneficiaries

- Measuring wellbeing at the start of a project, along with other important information on barriers or risks and demographics, gives useful information on beneficiaries and the targeting of the project.

### Measuring wellbeing lesson 2: measuring life satisfaction, pre and post

- Where a project seeks to make longer-term changes to the quality of life of participants,[30] it will often be relevant to measure life satisfaction before and after the intervention.

(29) With an average participant wellbeing of 6.3 at the start of the project, compared to a national average of around 7.6 for young people aged 10-17.

(30) For example, a project may seek to make longer-term changes to mental or physical health, to feelings of confidence, loneliness, belonging, or the opportunities which a person has for engaging in a certain activity.

# Background[31]

**Participation in sporting activities and exercise in natural environments is associated with improved wellbeing and decreased anger and depression.[32] People report feeling happier when in marine and coastal areas, and adolescents who spend time at the beach report feeling more relaxed and refreshed than in urban green spaces.[33]**

However, for children with long-term illness or disabilities, who have the same activity requirements as all children, there are often fewer opportunities to participate in sport.[34] In addition, young people living in urban areas, like London, can face significant barriers to accessing the coast and these barriers may be amplified for people with mental and/or physical health issues. For many young people, spaces and sporting activities are appreciated that afford them opportunities to develop their social identity and negotiate independence from adults.

Understanding the impact of surfing on physical and mental health is complex, but the importance of fostering a sense of belonging and identity is clear. Surfing is characterised by having its own language, rules, and values, all of which contribute to the development of identities.[35] Sociologists believe that it is the embodied performance or 'doing' surfing that contributes to young peoples' 'authentic' selves.[36] For disabled, urban children, 'lifestyle' sports like surfing provide an opportunity to become a 'surfer', to take risks, experience exhilaration, fear, pleasure, and fun.[37]

(31) This and the following section has been taken from Godfrey, C., Devine-Wright, H. (2022) Final report for Laureus, City to Sea (C2S) project,. Accessed Dec 10, 2024

(32) Levitas et al. (2007) : The Multi Dimensional Analysis of Social Exclusion, Thompson Coon et al., (2011) : Does participating in physical activity in natural environments have a greater effect on physical and mental wellbeing than physical activity indoors ? A systematic review

(33) White et al. (2013): Feelings of restoration from recent nature visits

(34) Moore, Clapham and Deeney (2018): Parents' Perspectives on Surf Therapy for Children with Disabilities

(35) Gleave and ColeHamilton (2012): A world without play: A literature review

(36) Wheaton & Beal (2003): "Keeping it real": subcultural media and the discourses of authenticity in alternative sport

(37) Bell, Phoenix, Lovell and Wheeler(2015): Seeking everyday wellbeing: The coast as a therapeutic landscape; Kelly (2018): 'I Need the Sea and the Sea Needs Me': Symbiotic coastal policy narratives for human wellbeing and sustainability in the UK

# The City to Sea project

**The Wave Project works with young people with mental health issues, social deprivation, or social isolation[38] and uses surfing and paddleboarding to reduce anxiety and increase confidence and social interaction. The Wave Project takes referrals from CAMHS and other statutory agencies.[39]**

The ethos of The Wave Project is one of positivity, acceptance, and fun. The 'surf therapy' provided by The Wave Project consists of six weekly group sessions that typically last two hours. Groups can accommodate up to 12 participants, with everyone receiving one-to-one support from a Wave Project surf volunteer. Participants engage in beach-based and water-based activities including beach games and surfing. Both the sea and sport provide a restorative and sensory experience that motivates young people to learn new skills, overcome challenges, and socialise. After completing the six-week course, participants can join a follow-on Wave Project 'Surf Club' that focuses on social connectedness, communication, and developing teamwork skills. This is presented in Figure 3. Former Wave Project participants are invited to volunteer as peer mentors on subsequent Surf Therapy courses.

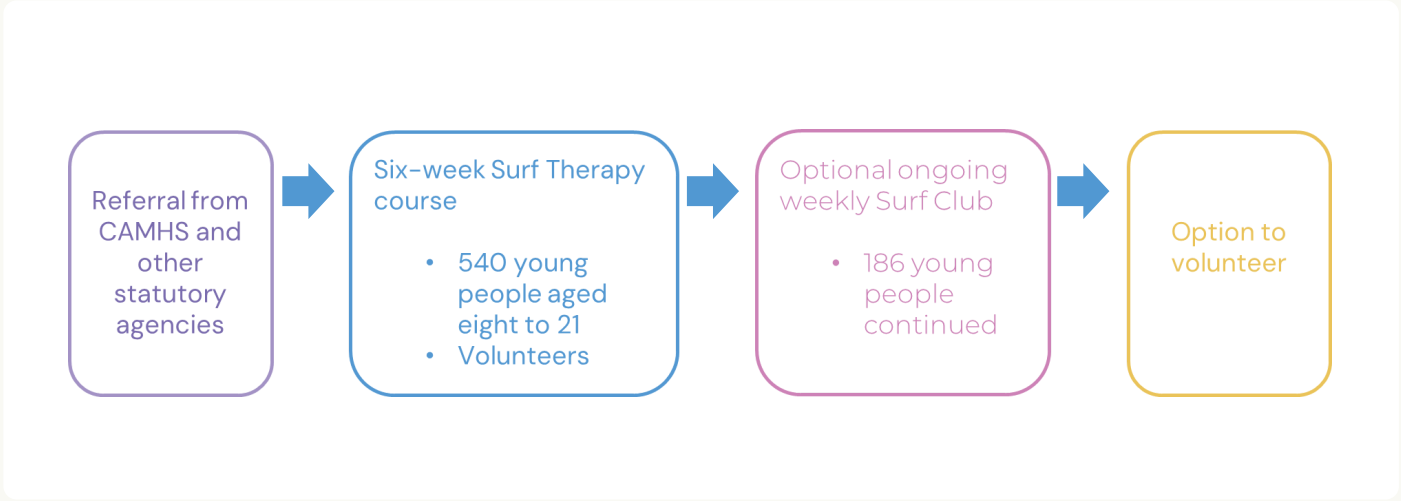
The City to Sea project extended The Wave Project to 540 at-risk young people aged eight to 21 from inner city London[40] and urban locations in Cornwall.

(38) Around a third of The Wave Project participants have diagnosed mental health issues, a third suffer social deprivation or family breakdown, and a third are socially isolated because of a disability, being a young carer, or being bullied. Referrals are taken from statutory agencies.

(39) The Wave Project intentionally targeted referrals of young people who were isolated and vulnerable because of complex needs such as mental health problems, disability, and illness, with a target of 40% of participants self-identifying as disabled. Some may have experienced trauma and crisis, or face problems exacerbated by self-harm, challenging behaviour, eating disorders, and/or other difficulties.

(40) Water-based sessions within London and surfing activities on the coast near Brighton (a journey of approximately two hours 15 minutes by private vehicle from Southwark).

Figure 3: The Wave Project journey



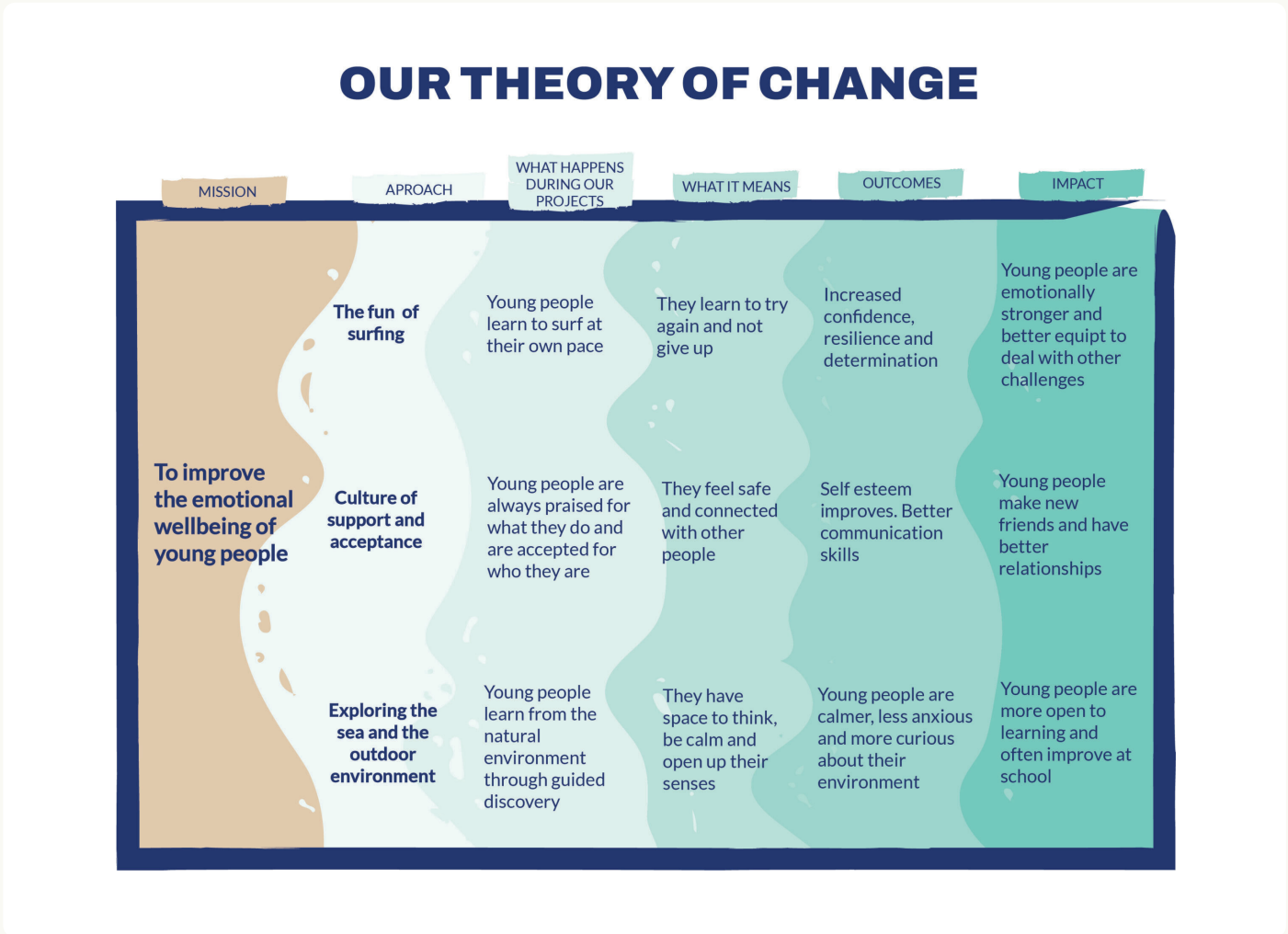

> "I feel that all my worries disappear when I'm on my bodyboard. It is the place I feel most relaxed"
>
> – City to Sea participant

These aims and the steps to achieve these changes are broadly reflected in the wider theory of change for The Wave Project (see Figure 4). The Wave Project is the overarching charity which delivered City to Sea alongside Laureus Sport for Good.

## The specific aims of the City to Sea project were:

1. Improved wellbeing, self-esteem, and resilience.
2. Reduced social isolation and better, stronger relationships.
3. Improved emotional development, soft skills, and improved life skills through volunteering.
4. Disabled people feel more able to participate in physical activity/sport.
5. Improved perceptions of disability.

Figure 4: The Wave Project theory of change



# Our approach

We used a break-even wellbeing analysis to understand whether the project offered value for money. We compared the wellbeing benefits to the costs to assess how much of the observed improvement in wellbeing would need to be attributed to City to Sea – and over what time period – for the benefits to outweigh the costs of the programme.

We followed a five-step methodology, described below.

## Step 1: estimate the average wellbeing benefit of Surf Therapy per participant

We used data on the change in life satisfaction of Surf Therapy participants, specifically data gathered two weeks before the Surf Therapy course and after the six-week course.[41]

This question was part of a survey administered to 543 clients across London and Cornwall and completed, at least partially, by 435 clients. Data was collected for 407 individuals at baseline and 329 individuals after the Surf Club ended.

Although this is a good response rate, with just over 60% of participants completing the post survey, it may not have been the same individuals who completed the pre and post survey, since data was not matched. We can therefore not attribute all of the change to City to Sea, since a proportion is likely due to changes outside of the programme, or due to differences in the wellbeing of those who completed the pre and post surveys.

(41) With post-interview from 0 to two weeks after the end of the course. For 34% of young people, their engagement continued beyond this point, through participation in weekly Surf Club sessions and for a small number, as volunteers for Surf Therapy.

To calculate the total potential wellbeing improvement, known as WELLBYs or wellbeing years, we need to multiply three key numbers together:

- The average improvement in life satisfaction[42] (0.78).
- The number of young people who completed Surf Therapy during the course of the City to Sea programme (540).
- The length of time for which the improvement in life satisfaction lasts for. We used a range of assumptions for this. For example, where the change was assumed to continue for six months, we use a factor of 0.5.[43] This figure was 0.25 where the change was assumed to continue for three months and 0.11 where the change was assumed to last only six weeks, the length of the course (and the minimum time between the pre and post surveys[44]).
- This means that if the effects of the programme persisted for three months, the total potential wellbeing impact of the programme would be 105 WELLBYs (0.78 * 540 * 0.25). However, it should be noted that we do not have evidence for how much of this improvement can be attributed to the programme itself.

(42) On a 0-10 scale.
(43) 6/12 months.
(44) Post surveys were completed between 0 and two weeks after the end of the course.

## Step 2: estimate the 'monetary value' of the wellbeing benefit of Surf Therapy per participant

The UK government has a recommended monetary value of £13k in 2019 prices for one WELLBY – a 1-point change in life satisfaction on a 0-10 scale[45] sustained for a year. This is the equivalent of £15,920 in 2024 prices. This monetary value is typically applied where there is confidence in causality. Since we cannot have full confidence in causality, the monetisation applied here and associated value for money scenarios can only be illustrative.

This means that if the effects of the programme persisted for six months, the total potential wellbeing impact of the programme would have a value of £1.7 million (105 *£15,920).

## Step 3: estimate the economic costs of the service per participant in consistent price years

We used financial data from Spirit of 2012 and The Wave Project of the total spend of the project over the three years from 2018 to 2021. This included grant funding from Spirit of 2012 and match funding from a range of additional donors, and included staff costs, overheads, evaluation, costs of equipment, surfing sessions, and even transport to the coast.

We expressed this in 2024 prices, using the GDP deflator,[46] to give a total cost of £0.46 million in 2024 prices. This figure is not the total amount of money which was spent by the charity. Rather, it represents the value of this spend, if it were to be spent 'today'. Due to inflation, £100 today can buy less than it could have bought in 2018. This means that £100 spent in 2018 has more value (it was able to buy more) than £100 today. All costs are expressed in the same 'price year' so they can be compared more equally, i.e. £1 has the same value no matter when it was spent.

We split this total figure in 2024 prices by the total number of Surf Therapy participants who completed the course over the three years – 540 – to give a cost per participant of £861 in 2024 prices.

(45) HMT (2021)

(46) ONS (2024): Gross Domestic Product at market prices: Implied deflator. Accessed March 31, 2025

## Step 4: carry out break-even analysis to illustrate the percentage of the wellbeing change which would need to be due to the project over specific timeframes for the project to break even

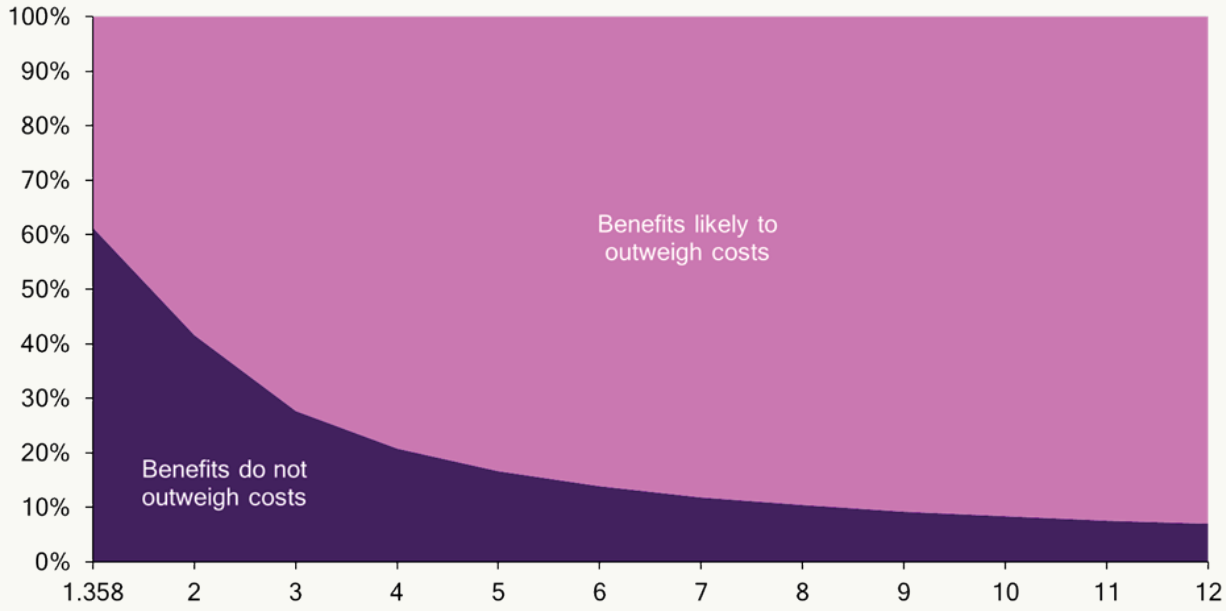
There are two key uncertainties which influence the value for money estimates: the length of time for which benefits last for and the proportion of the change in wellbeing which could be attributed to the project.

We used the above figure of the cost per participant, alongside the range of these two uncertainties, to create a 'heatmap' to demonstrate under which assumptions the project is likely to break even. The lowest timeframe is six weeks, which is the shortest timeframe between pre and post surveys.

We combined different scenarios of attribution and length of time of benefits to assess the 'break-even point', the point where the wellbeing benefits are equal to or greater than the costs. These scenarios are presented in Figure 5 below.

Figure 5: Across a range of scenarios, the wellbeing benefits of City to Sea are likely to outweigh costs

Proportion of improved wellbeing which would need to be attributed to City to Sea for benefits to be greater than the costs, if benefits persisted for 1.4 up to 12 months




Notes: Wellbeing was measured by life satisfaction on a 0-10 scale. The life satisfaction question was part of a survey administered to 543 clients across London and Cornwall and completed, at least partially, by 435 clients. Data was collected for 407 individuals at baseline and 329 individuals after the Surf Club ended (6 to 8 weeks following baseline).

Timeframe for length of time of wellbeing benefits starts at 1.36 months, which was the minimum length of time between pre and post surveys. The time period between 1.4 and 2months is shorter, although represented with equal space as a month in the diagram.

Source: PBE analysis of data from the Wave Project from 2018 to 2021

We considered a timeframe of three months to be a conservative but realistic timeframe for benefits. We also explored the potential for benefits lasting six months, in line with general research of the wellbeing benefits of social inclusion projects in the UK,[47] and as high as 12 months. A wider discussion on these uncertainties is set out in Box 5.

## What are realistic assumptions for duration and attribution?

What is a realistic assumption for the timeframe of benefits? The project specifically targets and found significant improvements in self-esteem, confidence, and connection to others: factors which can underpin changes in longer-term wellbeing.[48] This suggests that an ongoing change, beyond the six weeks of the course, could be likely.

Evidence from earlier Surf Therapy courses in the UK demonstrated significant increases in some measures of wellbeing that persisted up to three months after programme completion.[49] Even though life satisfaction was not specifically measured in this report, it supports the assumption that changes in certain aspects of wellbeing are maintained well beyond the end of the course.

What is a realistic assumption for attribution? The proportion which can be attributed to Surf Therapy rests upon two questions:

Firstly, what is the scale of the wellbeing improvement? The data isn't available for the same individuals which means it could be possible that changes in average wellbeing are partly due to different people (with different levels of wellbeing) answering, rather than a change in wellbeing for an individual. Twenty per cent of individuals did not complete the post interview.[50]

(47) See Frijters and Krekel(2021): A handbook for wellbeing policy-making in the UK: history, measurement, theory, implementation, and examples pg 173. One-off social inclusion programmes (e.g. a series of cooking classes) can raise life satisfaction by up to 0.5 points for six months (but should be expected to then tail off again).

(48) Also supported through case study evidence and examples.

(49) Devine-Wright, H. & Godfrey, C. (2020): The Wave Project: Evidencing Surf Therapy for Young People in the UK

(50) If an additional 20% did not complete the final interview (i.e. the difference between pre and post) and had 0 increase in wellbeing, this would reduce the change from 0.78 to 0.5 on average, or 60% of the average reported change.

Secondly, how much can be attributed to the programme as opposed to other changes occurring in people's lives? The wider context in the UK is one of declining average wellbeing of young people between 10 and15 over the time period in question. This is helpful context, but we cannot use it to draw conclusions for this sub-group over this timeframe, given that City to Sea participants were not representative of the UK averages, plus national averages are measured on a yearly basis and the data from City to Sea is a change over six weeks. Even the weather or season may have played a role in changes in reported wellbeing.(51)

Elsewhere, a 40% optimism bias reduction is recommended where there is uncertainty in benefits.(52)

This gives us no clear conclusion of an assumption for attribution, but around 60% could be considered a 'rule of thumb', consistent with analysis from elsewhere.

## Step 5: consider the role of volunteers

We estimated how the cost effectiveness (per participant) would change if the estimated wellbeing benefit to volunteers was incorporated. This value is not from the evaluation, but is applied based on causal evidence of the impact of volunteering on wellbeing. It was applied as an additional sensitivity.

We used the reported number of long-term volunteers for Surf Therapy plus Surf Club (535) and the total volunteer hours of 21,792 to estimate an average of 41 hours per volunteer. We applied estimates of the causal wellbeing benefit of (bi)weekly volunteering of 0.2 life satisfaction points(53) to roughly estimate the potential wellbeing benefit in addition to the change for participants. This should be considered a rough guideline of the type of additional benefit which could come from volunteers. We have low confidence in this estimate, since the results do not come from the project itself. We applied it only in sensitivity analysis.

Box 6 covers wider methodological points on accounting for volunteering in cost benefit analysis.

(51) The Children's Society (2023): The Good Childhood Report. Accessed November 10, 2024.

(52) For example, in the GMCA (2014) Cost Benefit Analysis Guidance for local authorities, table 7B. Accessed November 10, 2024.

(53) Krekel et al (2024): Happy to Help: Welfare Effects of a Nationwide Volunteering Programme. Accessed Dec 10, 2024. This is a Randomised Control Trial of the NHS responders scheme, which is considered a close enough proxy for the type of volunteering purpose.

## The costs and benefits of volunteering

In some evaluations/assessments, volunteering is included as a cost, since time is a resource which is limited and 'expended' in this way. In other assessments, volunteering time is included as a benefit, in terms of costs saved, since the organisation has saved the money which would have otherwise been required to pay staff. In other analysis, the wage rate is used as a proxy for the benefit, for the 'value' of the work which is done by the volunteer.

In this case, the benefit to recipients/the organisation is directly estimated through measuring the wellbeing benefits of participants in Surf Therapy, requiring no further proxy for social value. In addition, high-quality evidence in the UK demonstrates that volunteers gain wellbeing from their action, in spite of time being expended. This is to say that the estimates of wellbeing benefit from volunteering are in a sense a 'net' wellbeing benefit, from which the wellbeing 'cost' of using time has already been subtracted.

This means that only the wellbeing benefit to participants and the wellbeing benefit to volunteers need to be included to arrive at a full value of the wellbeing benefit, or social value.

## Key assumptions

- As described above, the wellbeing change measured between the start and end of the six-week Surf Therapy course may not all be due to participation in Surf Therapy. This is for two key reasons. Firstly, the scale of the wellbeing improvement is not clear. The data is not available for the same individuals before and after the course, so we could have a bias if those seeing the biggest improvemen`ts were more likely to remain for and answer the follow-up survey.(54) Secondly, there will have been other factors changing in the participant's lives over this six-week period, which may also have influenced their reported wellbeing (including changes in family, friends, school).(55) The scenarios explore how different assumptions of attribution would change the value for money implications.

- The duration of benefits is also uncertain, beyond the six weeks when wellbeing was measured post-Surf Therapy.

The impact of these key assumptions has been explored through a range of scenarios reported in our results.

(54) While there is insufficient data to assess how much of the benefits can be attributed to City to Sea, we know that it was not all of the same individuals who completed the pre and post questionnaire. 20% fewer individuals completed the post questionnaire compared to the pre questionnaire. The charity noted that parental engagement often influenced the likelihood of the post questionnaire being filled out.

(55) However, the short timeframe between pre and post responses means it is less likely that there are significant changes in some of the key factors which influence child wellbeing, such as mother's mental health, family income, and school teachers.

# Results

### Participant quote:

"P looked the happiest I've seen him in years. It's boosted his confidence and self-esteem. At school the teachers reported a noticeable rise in confidence, he's putting his hand up in class, and he stood up for himself against a bully. It's helped him grow and show him that he can achieve things he never thought possible"

### Parent of a surf therapy participant

### Participant quote:

"A enjoys being in the water and was happy for the opportunity to do this without her younger siblings as I am unable to accommodate this. A has become positive but needs to maintain this long term as it's often short lived. We have a long journey to make but Wave Project has been an amazing part of this, which I am truly grateful. It has given me a reason to get up and out on a Saturday and at least for one day a week the children have done something enjoyable"

### Parent of a surf therapy participant

The average life satisfaction score of beneficiaries improved by 0.8 points on a scale of 0-10 after the six-week Surf Therapy course.[56] Even though only 329 participants completed the follow-up survey, we applied this change to all 540 participants who completed the Surf Therapy course.[57] However, at present, we cannot be certain how much of this improvement is down to the programme (the attribution) and how long this improvement may have lasted for, beyond the six weeks of the course. We have used a range of assumptions to explore this:

- If all the wellbeing improvement could be attributed to The Wave Project and lasted for three months (i.e. for 1.5 months beyond the end of the course), this would deliver a potential economic benefit worth around £3,100 per person supported.[58] Compared to costs per participant of £860, this means that less than a third (28%) of the improvement in wellbeing experienced by participants needs to be down to the programme for the benefits of the programme to outweigh the cost.

- If these benefits lasted for three months and the programme was responsible for 60% of the benefits, then the programme would deliver just over £2 of benefits for every £1 spent.

- On the other hand, if the effects lasted longer, then the benefits could be far greater. For example, if they lasted for six months and 60% of the improvement in wellbeing could be attributed to The Wave Project, then the programme would deliver over £4 of benefits for every £1 spent.[59]

(56) When measured before the course and 0 to two weeks after the six-week course.

(57) View from the charity that completion of follow-up response was unrelated to the perceived benefit or level of wellbeing change. A completed follow-up response was more often related to level of parental involvement.

(58) £1,410 if the benefits only lasted for six weeks and £6,210 if the benefits lasted for six months.

(59) If the benefits lasted for six months, then only 14% of the benefits would have to be attributable to Surf Therapy for the benefits to outweigh the costs.

The estimated benefits would be even higher if we include the additional wellbeing benefit to volunteers. Since the wellbeing of volunteers was not measured, we used the best causal evidence of the impact of volunteering on wellbeing[60] to calculate the total assumed wellbeing gained by volunteers[61] and divided this by the number of participants to estimate the additional wellbeing of volunteers per participant. This increases the assumed wellbeing benefit across the project and on average per participant and therefore reduces the percentage attribution required for the project to break even. Specifically:

- If the benefits for participants lasted for three months, only 23% (compared to 28%) would have to be due to City to Sea for the project to break even.

- If 60% of the improvement in wellbeing could be attributed to City to Sea and the participant wellbeing lasted for six months, then adding the wellbeing benefits to volunteers would mean the programme would deliver £5.10 of benefit for every £1 spent (compared to £4.30 without).

(60) Krekel et al. (2024). Change of 0.2 life satisfaction points for weekly volunteering.

(61) We estimate this based on the total number of volunteer hours, which is split by total reported volunteer hours, to come to an average number of hours per volunteer. Since each weekly session required three hours of volunteering, this would represent 13 weeks in total, or three months of weekly volunteering, which is what the 0.2 WELLBY benefit is based on. In reality, this three months of benefit could be split across a year on average, given that volunteering is seasonal and volunteers sign up for a six-week course, for example in the autumn, then volunteer again for another six-week course in the spring. We assume that the wellbeing benefit is experienced only during the weeks of volunteering, which is likely to be an underestimate, given the positive reports from volunteers of an increase in feelings of community and friendships. This continues over winter when there are wider Surf Club activities and socials.

We also considered what would happen if only the participants who completed the post survey experienced benefits. This lowers the benefits, although it still appears plausible that the programme offers value for money. In this situation, if the benefits lasted for three months, 46% (compared to 28%) would have to be due to City to Sea for the project to break even. If benefits lasted for six months, 23% would have to be due to City to Sea (compared to 14%) for the project to break even.

## Conclusions

**Overall, despite the limitations in the evidence available at the time, our analysis suggests that it's plausible that City to Sea offered good value for money across a range of scenarios. The case is further strengthened by including other potential wellbeing benefits such as volunteering.**

This case study demonstrates how programme wellbeing value for money can be estimated, even when there is not all the quantitative information available to be certain of the average monetised wellbeing impact. A nuanced picture of how the project may have changed individuals' lives can be broadened with qualitative data and case studies.

A mum describes the impact of The Wave Project on her son, whose dad had recently passed away:

"He has always been active, Sea Scouts, kayaking, a real 'normal' lad, but the grief just hit him. He went missing one night and I knew then how desperate he was. The Wave Project gave him a reason to keep living, to keep going. After that first session he got in the car with me and his brother and he grinned and his brother said to me, 'What did they do to him?' I said, 'Took him surfing.' He said, 'No, I mean, what did they give him?' [He] had found what he needed to be happy"

# City to Sea technical annexes

## Annex A1 – Data required for break-even scenario analysis

Data which was required for this scenario analysis:

**We used the following data collected by The Wave Project from 2019 to 2021[62] and reported to Spirit of 2012:**

- Reported life satisfaction on a 0-10 scale before participation in programme.
- Reported life satisfaction on a 0-10 scale following six-week programme.
- Numbers of overall participants and number of completed life satisfaction responses.
- Average time between 'pre' and 'post' surveys.
- Good theoretical evidence of why this project would make a difference, backed up with evidence from elsewhere. The results from further wellbeing questions from this survey were referred to, to understand the potential pathways to the final wellbeing impact.
- Financial reporting of grant funding and match funding, split by calendar (or financial) year.
- For additional scenario with volunteers, number of volunteers and volunteer hours with information on engagement.

(62) The initial six months of funding from October 2018 were used to establish the project.

Further wellbeing questions were asked by The Wave Project, to understand in what ways wellbeing was changing. The City to Sea evaluation[63] has full detail of the questions asked and results from the pre/post analysis.

## The Wave Project is currently following these steps with its evaluation:

- Matched pre and post data (unique identification number).
- Gathering data for different types of engagement: e.g. participants in Surf Therapy, participants in Surf Club, and volunteer programme.
- Longer-term follow-up, including for those no longer participating.
- If The Wave Project would continue to measure life satisfaction pre and post the intervention, it would be possible for them to estimate a wellbeing benefit cost ratio for their work.

(63) Godfrey, C., Devine-Wright, H. (2022) Final report for Laureus, City to Sea (C2S) project,. Accessed Dec 10, 2024


# Annex B: EmpowHER case study

## Summary

EmpowHER was created in response to an open call from Spirit of 2012 and the #iwill Fund for a social action project inspired by the centenary of women's suffrage in 2018. The programme provided young women and girls (YW&G) with opportunities to make change in their communities. UK Youth led the programme between 2018 and 2021, in partnership with the British Red Cross (BRC) and Young Women's Trust (YWT). It received £2.4 million of combined funding from Spirit of 2012 and the #iwill Fund,[64] with 1,880 YW&G successfully completing the programme over this time period. This report assesses the wellbeing value for money of the EmpowHER programme, as delivered between 2018 and 2021.

UK Youth gathered matched data[65] for all participants at the beginning and end of participation in EmpowHER, including demographic, wellbeing, and participation data. This enabled PBE to create a 'matched group', a group with similar characteristics, to explore how wellbeing could have changed over the time period in the absence of the programme.

It is a useful case study of the conclusions which can be drawn with this further data, which enables greater confidence in the change which was due to the programme. The analysis gives a benefit cost ratio: a comparison of the monetised wellbeing benefits to the monetised financial costs.

(64) The #iwill Fund is made possible thanks to a £66 million joint investment from The National Lottery Community Fund and the Department for Culture, Media and Sport to support young people to access high quality social action opportunities.

(65) For instance, for the same (anonymised) respondent there was data on the life satisfaction score before participation (the date of this response), after participation (the date of this response), as well as a range of further demographic characteristics such as ethnicity and age.

## We found that:

- YW&G completing EmpowHER experienced an average improvement of 0.9 points in life satisfaction over an average of 4.4 months. Since the wellbeing of non-participants could have decreased by around 0.4 in this same period,(66) this represents a total of 1.3 life satisfaction points over 4.4 months.
- On this basis, we estimate the value of these wellbeing benefits to be around £6,500 per participant, based on HMT's approach to valuing wellbeing benefits.(67)
- Given the programme cost on average £1,300 per participant, it is estimated to generate around £5 wellbeing benefits per £1 spent.
- This conclusion – that the programme offers good value for money – is valid across a range of alternative assumptions.

(66) Analysis was carried out for the change over a year, an estimated 0.8 drop over the year.

(67) Where the difference in life satisfaction is multiplied by the relevant number of months and the HMT value. Note that these figures are based on all participants: 91% of participants completed EmpowHER. For completed participants, the costs and wellbeing benefits are £1,473 and £7,562 respectively.

Our analysis demonstrates the value of having matched data – and the additional understanding about a programme when it is possible to estimate what could otherwise have happened, in the absence of the intervention. It is also a good example of the value of assessment and learning. UK Youth was keen to build evidence on the wellbeing benefits of social action interventions, particularly when focused on YW&G with lower starting wellbeing. Information on participation and wellbeing benefits were able to help the charity to steer their intervention for future cohorts.

## Measuring wellbeing lesson 3: matched data

- **Where life satisfaction is collected pre and post, there is additional value in recording this data at the individual level**, so there are records of each individual's demographics and their change in wellbeing. This 'matched' data gives us more confidence than aggregates, especially when different participants may have answered the pre and post surveys.
- In addition, matched data gives the option to create a comparison group. This allows us to estimate what could have happened in the absence of the project and gives us even further confidence that the change is due to the project itself.

## Measuring wellbeing lesson 4: test, learn, and adapt

- UK Youth specifically planned in learning pauses between cohorts of the EmpowHER programme, to assess the findings and amend delivery based on this. Due to this approach and a critical assessment that there should be a higher participation from YW&G with specific barriers, they amended the process for targeting and were better able to include young people with specific barriers in later cohorts.[68] Perhaps partially as a result, the pre-post change in life satisfaction was higher in the following cohorts.

(68) See earlier discussion, 'Reaching those with greatest need'.

> "It makes you feel excited, like you're making a change and the world will one day be better because you're doing something about it"
> – Young person

# Background[69]

**Children and young people's wellbeing is in decline and has been for some time.[70] YW&G, especially those from disadvantaged backgrounds, are even more likely to experience poor wellbeing.[71]**

At the time of evaluating EmpowHER, UK Youth identified that girls were more prone to concerns about how they look, had a much higher likelihood of self-harming, and a higher likelihood of having (or developing) an eating disorder.[72] A survey by Girlguiding pointed to some potential causes of these discrepancies.[73] Almost all the girls surveyed (95%) said that the advertising industry should show more positive, diverse representations of women. Many felt angry about gender stereotypes in the media and the products sold to them, but despite this awareness many girls internalise these negative messages. More than half of the girls surveyed (55%) said that prevalent gender stereotypes affect their ability to say what they think, and they feel that stepping outside the norm will result in bullying and teasing from their peers. In addition, girls who identify as Black, Asian, orethnic minority, LGBTQI+, or disabled face additional barriers when it comes to feeling safe and confident to be themselves.[74]

(69) Content from UK Youth and Spirit of 2012 (2022): EmpowHER Evaluation Report from 2018-2021. Accessed October 15, 2024
(70) NHS (2023): Mental Health of Children and Young People in England, 2023 – wave 4 follow up to the 2017 survey. Accessed October 15, 2024
(71) The Children's Society (2023) The Good Childhood Report. Accessed October 15, 2024
(72) Brooks et al. (2020) HBSC England National Report: Findings from the 2018 HBSC study for England. Accessed Feb 12, 2025
(73) Girlguiding (2017) Girls' Attitude Survey 2017. Accessed Feb 12, 2025
(74) Girlguiding (2020) Girl's Attitude Survey 2020. Accessed Feb 12, 2025

However, engagement of young people in social action provides an opportunity to boost young people's wellbeing. Evidence suggests that inclusive and meaningful social action opportunities[75] and the influence of role models can improve the wellbeing of YW&G and inspire them to lead change in their communities. In addition, YW&G point to youth groups as places where they can be themselves, learn new things, and build positive and supportive relationships.[76] A trusted support network is a key mitigating measure for poor wellbeing.[77] A Cabinet Office report found positive improvements in resilience, problem solving, empathy, and community investment among YW&G who participated in social action.[78] Finally, the National Youth Social Action survey (2018) (funded by DCMS) found that volunteering and social action was particularly helpful for YW&G since – once involved – girls are more likely to believe their actions can have an impact on their community.[79] Yet YW&G with low socioeconomic background may have been excluded from social action opportunities due to lack of access or the necessary skills or confidence to engage.

This evidence prompted Spirit of 2012 – with the #iwill Fund – to create an open call for funding applications for girls at key transitional ages. Spirit of 2012 wanted to test whether social action could be a route to increasing girls' wellbeing. UK Youth, BRC, and YWT responded to this call, developing the EmpowHER programme.

(75) Youth social action refers to activities that you can do to make a positive difference to others or the environment. There are lots of ways in which you can take practical action to make a positive difference. It can take place in a range of contexts and can mean formal or informal activities. These include volunteering, fundraising, campaigning, or supporting peers.

(76) Girlguiding (2020)

(77) Brooks et al. (2020)

(78) Kirkman et al. (2016) Evaluating Youth Social Action – Final Report. Accessed Feb 12, 2025

(79) Knibbs et al. (2019) National Youth Social Action Survey 2018. Accessed Feb 12, 2025

# The EmpowHER project

**EmpowHER was designed to address low wellbeing among YW&G through supporting their involvement in inclusive, meaningful social action. Co-designed with young people, the programme supported YW&G through a journey of social learning. It provided activities and learning around individuals' rights, wellbeing, empowerment, and resilience, providing a foundation from which to partake in (and ultimately lead) social action opportunities. The programme was targeted at 'unusual suspects' – YW&G between the ages of 10 and 20, with low wellbeing, and who may have been excluded from social action opportunities due to lack of access or the necessary skills or confidence to engage.**

Four cohorts of YW&G took part in EmpowHER between 2018 and 2021. UK Youth led the programme, in partnership with BRC and YWT. The three partners brought together their expertise in youth work, social action, campaigning, national programmes, and gender specific advocacy, and worked closely on the planning and delivery of EmpowHER across the country.

With an open network of over 9,000 youth organisations and nation partners, UK Youth is focused on unlocking youth work as the catalyst of change. From across UK Youth's network, 21 delivery partners were recruited and trained on the aims, structure, and delivery requirements of EmpowHER, and worked closely with UK Youth to flexibly adapt elements of the programme to their local context. The youth organisations and youth workers brought value to the programme through their knowledge and expertise in their local context, including in engaging 'unusual suspects' and adapting the programme to meet the different needs of young people.

Over the six months of each cohort, the YW&G met with a youth worker in small groups across the South West, West Midlands, North West, and East of England for at least 12 weeks, completing a minimum of 30 hours.

The programme, as delivered between 2018 and 2021, had three key elements: social learning, facilitated by a youth worker; social action, led by the YW&G; and social leadership, with opportunities for the YW&G to further develop their confidence and lead change. Underpinning these elements was regional working, sharing of best practice, and celebrating and recognising YW&G achievements. This is set out in Figure 6. Figure 7 summarises the inputs, activities, outputs, and outcomes contributing to the final change in wellbeing.

Figure 6: The EmpowHER model

Figure 7: Simplified logic model for EmpowHER

| **Inputs** | **Activities** | **Outputs** | **Outcomes for EmpowHER participants** | **Impacts** |
|---|---|---|---|---|
| Youth workers trained in EmpowHER model | **Social learning** sessions: learning around individuals' rights, wellbeing, empowerment, and resilience | Number of young women and girls supported | Reduced limiting perceptions of themselves and their gender | Improvement in wellbeing |
| EmpowHER overheads | Involvement in inclusive, meaningful **social action** | Number of social action projects | Improved perceptions of social cohesions | |
| Access fund to reduce barriers associated with certain groups participating | **Social leadership** | | Improved ability to lead and identify change | |
| | Mix of fun and serious activities | | Development of social action habits | |
| | Exposure to female role models | | | |

## EmpowHER case study (80)

- Following the murder of Sarah Everard, the young women at InUnity wanted to address negative behaviours of the young men in their year group. They created posters containing tips about how their male peers can support the women in their year group to feel safe. They also organised an expert to facilitate a workshop to address negative language towards young women and young LGBTQI+ people, encouraging discussion and exploring questions such as "why do you think you can comment on our bodies?" The young women have been working alongside their senior leadership team to address their school PE uniform.

## EmpowHER case study

"The thing that motivated me to start EmpowHER: because I am Asian and we live in a very cultural community also… because it was at a local community centre I was able to go, and honestly if I didn't start EmpowHER I would have just stayed at home watching Netflix with a bag of crisps"

– Participant, Uniting Communities Organisation

(80) Source: UK Youth and Spirit of 2012 (2022): EmpowHER Evaluation Report from 2018-2021. Accessed October 15, 2024

Before EmpowHER, one particular participant had not been involved in social action. She joined the Uniting Communities Organisation (UCO) EmpowHER group where she and her group members participated in first aid training delivered by the British Red Cross and went on to organise a first aid community day where they shared what they had learnt. Approximately 45 people from their local community attended and learnt a range of first aid techniques. The young women and girls realised that some of the attendees didn't speak English, so they provided translations in Punjabi to ensure everyone could access the information and benefit from the training.

This experience has given the young women and girls the confidence to speak publicly and they feel empowered by the first aid skills they acquired.

> The participant mentioned: "If someone is dying now I know what to do, I can go get a defibrillator, I need to go and help people, I can't be shy all my life, I need to make myself useful."

Today, she is a volunteer with UCO, supporting the young women and girls in their current EmpowHER groups. She helps to boost the young people's confidence and self-esteem and gets involved in the young people's social action projects. She has also gone on to share her experience publicly at the British Red Cross Volunteer Mobilisation conference and UK Youth's Movement webinar. EmpowHER has had an impact, not just on this participant, but on all the young women and girls she has volunteered with, and the conference and webinar attendees who have been impressed by her confidence and passion for the project.

# Our approach

We linked the wellbeing data (life satisfaction) collected by UK Youth through to the wellbeing value for money using a six-step process.

## Step 1: estimate the average change in life satisfaction in the young people who completed the EmpowHER programme

Addressing low wellbeing as experienced by young people themselves was one of the main aims of the EmpowHER programme. Spirit of 2012 required that all funded projects collected consistent data on wellbeing, including life satisfaction. For this report, we focus on the life satisfaction measure,[81] a single measure which is widely used and is monetisable for value for money analysis according to UK government guidance. UK Youth's evaluation of EmpowHER collected data for participating YW&G's life satisfaction at baseline (before the programme) and after completing the programme (a minimum of 30 hours and a minimum of 12 weeks),[82] up to six+ months later. This data was only gathered for cohorts 1, 2, and 4, since cohort 3 was impacted by Covid. This data for cohorts 1, 2, and 4 was used to estimate the average change in life satisfaction.

(81) Specifically, the life satisfaction question as worded by the ONS in population surveys: "For each of these questions I'd like you to give an answer on a scale of 0-10, where 0 is 'not at all' and 10 is 'completely'. Overall, how satisfied are you with your life nowadays?"

(82) Average completion rate of 91%, where completion was defined by EmpowHER as engaging in 30 hours of sessions over a minimum of 12 weeks, with a minimum of 10 hours of social learning and 10 hours of social action.

## Step 2: estimate what change in wellbeing may have taken place without EmpowHER

Here we aim to understand how wellbeing may have changed for participants if they didn't take part in EmpowHER, ultimately to help us understand the full impact of EmpowHER in the following steps. We estimated the change in wellbeing of a matched comparison group from the Understanding Society Youth survey, a panel survey.

We used the following information on participants to find a matched group in the Understanding Society Youth Panel:

- **Sex.**
- **Age.**
- **Ethnicity.**
- **Starting wellbeing score.**
- **Calendar year.**
- **Existence of a 'barrier'.**[83]

This analysis estimated that the wellbeing of a similar group fell in the following calendar year.

(83) Barriers were Y/N answers for the following headings (brackets give the data used for matching when barrier was Y):
- Low income (assumed to be lower quartile of household income).
- Poor mental health (SDQ Total Difficulties score greater than 14).
- Low education (score 5, 6, or 7 on "How do you feel about your school work?").

## Step 3: estimate the change in wellbeing due to EmpowHER

By comparing the difference in wellbeing between the matched group and the group who completed the EmpowHER programme, we can estimate the wellbeing impact of participation in the programme. Since the wellbeing of the matched control group fell while the wellbeing of EmpowHER participants increased, it suggests that EmpowHER may not only improve wellbeing, but also have protective characteristics, halting a drop in wellbeing which is observed in the teenage years.

The results from the matching are shown in Table 1, with the chosen approach highlighted in blue.

Table 1: Options and results from matching

| Engagement period with EmpowHER | n | Matching method[84] | Match on barriers | Average engagement length in days (months) | Change treatment | Change control[85] | Total impact on LS |
|---|---|---|---|---|---|---|---|
| 12 weeks+ | 417 | Propensity score | N | 133 (4.4) | 0.96 | -0.43 | 1.30 |
| 12 weeks+ | 476 | Mahalanobis | N | 133 (4.4) | 0.95 | -0.35 | 1.29 |
| 12 weeks+ | 417 | Propensity score | Y | 133 (4.4) | 1.17 | -0.43 | 1.59 |
| 12 weeks+ | 476 | Mahalanobis | Y | 133 (4.4) | 0.95 | -0.36 | 1.30 |
| 3 months+ | 382 | Propensity score | N | 145 (4.8) | 0.93 | -0.45 | 1.38 |
| 3 months+ | 388 | Mahalanobis | N | 145 (4.8) | 0.90 | -0.39 | 1.29 |
| 4 months+ | 272 | Propensity score | N | 164 (5.5) | 0.9 | -0.49 | 1.4 |
| 4 months+ | 275 | Mahalanobis | N | 164 (5.5) | 0.88 | -0.51 | 1.39 |
| 4 months+ | 272 | Propensity score | Y | 164 (5.5) | 1.04 | -0.52 | 1.56 |
| 4 months+ | 275 | Mahalanobis | Y | 164 (5.5) | 0.88 | -0.38 | 1.26 |
| 5 months+ | 185 | Propensity score | N | 184 (6.1) | 1.08 | -0.48 | 1.56 |
| 5 months+ | 185 | Mahalanobis | N | 184 (6.1) | 1.08 | -0.53 | 1.61 |
| 6 months+ | 76 | Propensity score | N | 264 (8.8) | 1.11 | -0.94 | 2.05 |
| 6 months+ | 76 | Mahalanobis | N | 264 (8.8) | 1.11 | -1.06 | 2.16 |

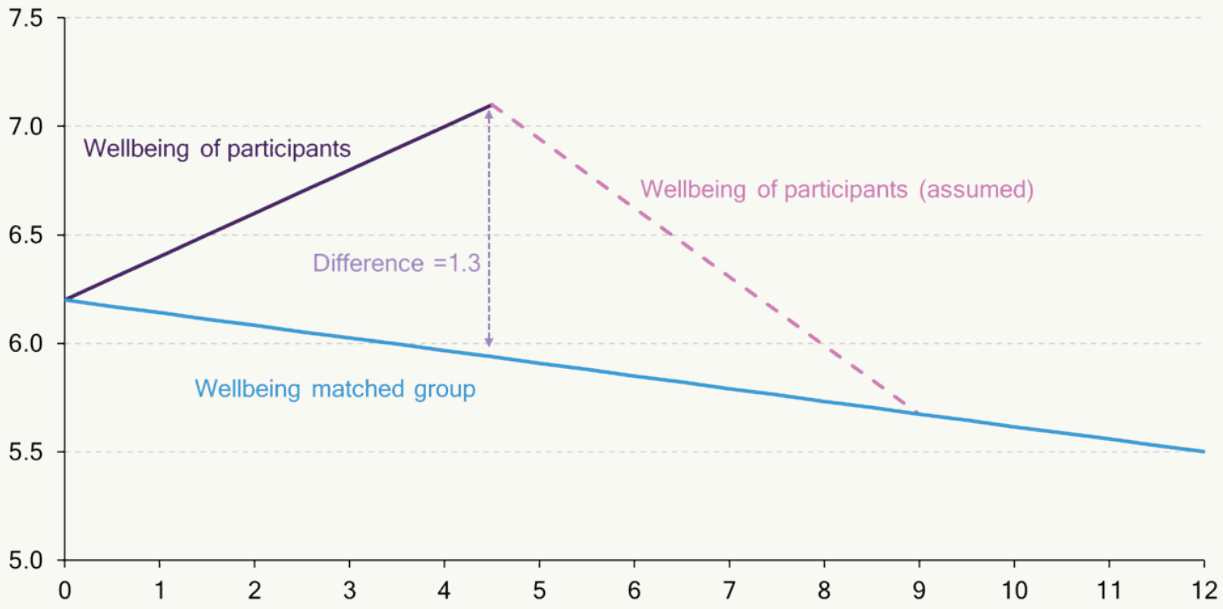
(84) Mahalanobis distance matching (MDM) and propensity score matching (PSM) are methods of doing the same thing, which is to find a group of units similar to treated units to arrive at a balanced sample (i.e. where the distribution of covariates is the same in both groups). MDM aims to match covariates as closely as possible, whereas PSM weights based on how well the covariates predict participation. In this analysis, both methods gave fairly consistent results, with the PSM method giving slightly bigger effect sizes.

(85) Where the change in the counterfactual has been proportioned by the average number of days.

We assessed the impact over 4.4 months, since this was the average number of months of participation for those who engaged for 12 weeks or more. As illustrated in Figure 8, this total change comes from i.) the average observed improvement in wellbeing over 4.4 months of 0.95 (0.34 pro rata, i.e. in wellbeing 'life years', or WELLBYs) and ii.) the estimation that, without participation, wellbeing would have otherwise dropped in the same time period. This was estimated by creating a matched comparison group. Specifically, the wellbeing of a matched comparison group fell by around 0.98 life satisfaction points over the whole year, which is estimated at ~0.36 over 4.4 months (the period in question).

Figure 8: Wellbeing of EmpowHER participants increased while wellbeing of the matched group was estimated to fall

Estimated average change in life satisfaction (on a 0-10 scale) for EmpowHER participants compared to a matched group over 0 to 12 months




Notes: Average change in life satisfaction (on a 0-10 scale) for EmpowHER participants is based on data points at 0 months and 4.4 months, with an assumed linear increase between these points and a linear decrease following the average 'post' measurement at 4.4 months. The estimate for 1,559 completed participants is based on responses from the 476 young people with full data to enable a reasonable match.

Matched group change in wellbeing is based on Understanding Society Youth panel data and estimated over 12 months.

Source: PBE analysis of data from UK Youth gathered between 2018 and 2021 and Understanding Society data. Further details in text To calculate the number of WELLBYs, which are a 1-point change in life satisfaction (0-10 scale) per person per year, the figure was apportioned by the number of months this is applied for.

To calculate the number of WELLBYs, which are a 1-point change in life satisfaction (0-10 scale) per person per year, the figure was apportioned by the number of months this is applied for.

For example, a change of 1 point for six months is 1 * 6/12 = 0.5 WELLBYs, since it only lasts for half a year.

### Calculation:

(Change in treatment group - change in control group) * length of days/days in a year

= (0.95 -0.36) * 133/365
= 1.30 * 4.4/12
= 0.48 WELLBYs

We multiplied this by the number of participants we assumed this benefit applied to, including all those who engaged for at least 12 weeks,[86] excluding the cohort of participants who engaged during Covid.[87]

## Step 4: monetise this change in wellbeing

We multiplied this change in wellbeing by the monetised value of a 1-point change in life satisfaction, commonly referred to as a WELLBY.

The value of a WELLBY comes from the HMT Green Book Supplementary Guidance on Wellbeing.[88] These values are uprated to 2024 prices, meaning that a single WELLBY has a central estimate of £15,900 (in 2024 prices).[89]
By monetising wellbeing, it becomes possible to integrate it seamlessly into economic evaluations, giving it equal footing with financial and productivity metrics.

Incorporating the 91% completion rate leads to a benefit per participant of £6,538 (compared to a benefit of £7,562 per completed participant).

(86) Some sites delivered the 30 hours over a more intensive period of time. Since this did not fit the EmpowHER model, which was designed to be delivered over a minimum of 12 weeks, the results from these groups were excluded.

(87) In a sense, assume that the benefits were no greater than 0. We assume that this group did not receive the same wellbeing benefit since the conditions and necessary delivery was different from the other cohorts and no wellbeing data was gathered during this cohort, to enable a comparison or to validate that the impacts were the same.

(88) HMT (2021) £10k-£16k in 2019 prices, with a central estimate of £13k. Associated discussion paper on monetising life satisfaction provides more background information on the derivation of the values.

(89) Rounded to nearest £100.

## Step 5: estimate the costs of the programme

UK Youth reported spend over the three years, including staff costs, overheads, further resources such as spend on celebrations or supporting YW&G with participation, and wider costs such as evaluation. This includes the costs for all participants, not just those who completed the programme. We excluded the costs for cohort 3 delivery and cohort 3 learning, which was the cohort during Covid, where delivery was interrupted and wellbeing was not measured. The rest of the costs, including overall programme development and overall programme learning costs, are included to estimate the total cost of the programme.

We expressed these total costs in 2024 prices, using GDP deflator figures.[90] This gave a total programme cost of £2.9 million in 2024 prices[91] and £2.3 million when only including cohorts 1, 2, and 4.

To estimate the average cost per participant who completed the course in these cohorts, we divided the total costs over the three years by the total completed participants.

## Step 6: assess the value for money

We compared the costs of delivery to the wellbeing benefits monetised in Step 4. This gave us an estimate of the wellbeing benefits produced for every £1 spent, an assessment of the value for money. We applied sensitivity analysis around this figure.

(90) ONS (2024): Gross Domestic Product at market prices: Implied deflator. Accessed March 31, 2025

(91) £2.4 million in nominal terms; £1.9 million excluding the Covid cohort.

## Key assumptions

- The matching rests upon the assumption that the evaluative wellbeing question asked in the Understanding Society Youth (USocY) survey is the same as the ONS4 life satisfaction question which was asked by UK Youth. Both are evaluative wellbeing questions, but use different scales with different labels. Specifically, the USocY survey asks, "How do you feel about life overall?" with labels of "happy" and "not at all happy" on a 1-7 scale, whereas the ONS question asks, "Overall, how satisfied are you with your life nowadays?" on a 0-10 scale. This assumption is consistent with wider academic research. The adult Understanding Society (USoc) survey, with labels of "completely satisfied" and "not at all satisfied" on a 1-7 scale, is considered consistent with the 0-10 ONS life satisfaction question. Initial analysis,[92] suggests that there is no change in the trend of responses when respondents switch from the youth survey at age 15 to the adult survey at age 16, therefore the labels and associated responses are considered sufficiently similar between the youth and adult survey. This enables us to assume that the responses from the matched panel, created from USocY, can be compared with the life satisfaction responses from EmpowHER.

- We assumed a non-linear translation from the 7-point youth scale to the 11-point adult ONS life satisfaction scale, based on aligning the cumulative distributions for all adults in 2021 between APS (11-point ONS measure) and USoc. For this non-linear translation, we assumed that the cumulative distribution for adults is the same as for children.[93]

- Data for the matched control group is only available at yearly intervals. However, to compare the two groups, it was necessary to estimate what change may have happened at 4.4 months. We assumed that the change over one year for the comparison group was linear and that the change at month 4.4 was equivalent to 4.4/12ths of the change observed for the whole year.

(92) Parkes (2025) The C-WELLBY: Towards a Universal Measure of Children's Wellbeing for Policy Analysis. Accessed April 15, 2025

(93) Further runs of the matching found that a linear translation resulted in a higher value, meaning that the assumption taken is more conservative.

- We assumed that the data gathered for the 476 young people with full data to enable a reasonable match[94] was representative of the 1,559 YW&G who completed EmpowHER in cohorts 1, 2, and 4.[95] A challenge with data coming from only a proportion of participants means that it could be unrepresentative of the whole group. They may have different characteristics and experiences that mean our estimates of benefits are not accurate. For example, those who had a larger wellbeing improvement from the project could have been more likely to fill in the questionnaires, which would bias the results upwards. However, the greater proportion of excluded participants came from other reasons, such as being outside the age range for which a match was possible.[96]

- The average length of participation **for those engaging for at least 12 weeks** was 133 days, ~4.4 months. We assumed that, on average, wellbeing falls again after this point (see Figure 8 ). The data from Table 1 shows that wellbeing benefits are higher for those who participate for five months or longer and six months and longer. Taking the figure for this group would overestimate the average impact, given that the minimum length of time for completion was only 12 weeks and 4.4 months was the average length of involvement for those engaged 12 weeks or more. We tested different scenarios for the duration of impacts in our sensitivity analysis.

- There could have been changes outside of the programme which influenced the wellbeing of participants. Even though a matched sample was created, which aims to 'control' for these outside influences which could have impacted the participants and non-participants alike, there may be factors relating to the programme. These caveats add some uncertainty, which were not possible to minimise further given the evidence available. Again, the robustness of our core results to variations in these assumptions were explored as part of the sensitivity analysis.

(94) The original data included responses from 1,114 individuals – over 70% of the completed participants. 35% of responses had to be dropped since they were outside the age range of 10-14, which would enable the matched comparison with the USocY survey. Since average wellbeing tends to drop for every year during the pre-teen and teenage years, exclusion based on age is unlikely to influence the change in wellbeing in a biased way. 20% of the remaining were dropped since they were missing a valid pre or post life satisfaction measure; 1% since the engagement year was not valid. The remaining observations were dropped for an engagement of less than 12 weeks.

(95) 1,880 completed participants minus 239 participants who completed EmpowHER during Covid (this group has been excluded since wellbeing data was not collected for this cohort), minus 5% who completed the required 30 hours but in less than 12 weeks.

(96) Only those within the starting age range of 10-14 were included, since the USocY survey was used, which collects data from 10 to 15-year-olds. A starting age of 15 would have meant that the participant would be outside of this age range by the following year, making a match not possible.

- The full costs of EmpowHER[97] may be underestimated if there were additional resources, such as venues, expertise, or time provided free of charge. Although these do not show up on the financial spend for the programme; they are 'real' costs which should be included in a full value for money estimate. On the other hand, the costs of future delivery of EmpowHER would be lower, since the initial fixed costs associated with developing the EmpowHER model would not be required for a future roll out. We explored scenarios with higher and lower costs in the sensitivity analysis, to test the robustness of our value for money estimates.

- Lastly, estimates may be conservative, because there could be potential cost savings to the health, education, and social security systems from improved wellbeing outcomes, such as reduced mental health costs, improved school attendance, improved employment chances, or even a reduction in criminal costs, that are not captured in our analysis due to insufficient evidence.[98]

# Results

YW&G completing EmpowHER experienced an average 0.95-point improvement in life satisfaction after an average of 4.4 months using data gathered between 2018 and 2021.[99] Since the wellbeing of non-participants could have decreased by around 0.36 in this same period,[100] this represents a total of 1.3 life satisfaction points over 4.4 months.

(97) As delivered between 2018 and 2021.

(98) In the #Untapped report, UK Youth estimates the benefits of youth work in general and the potential taxpayer savings. See Frontier Economics (2022): The economic value of youth work. Accessed March 24, 2025

(99) The central estimate rests on the assumption that wellbeing benefits do not last for longer than 4.4 months, the average timeframe at which the change in wellbeing was measured.

(100) Analysis was carried out for the change over a year, an estimated 0.8 drop over the year.

> "I learnt the importance of getting involved in the community and that if you're passionate about something, the impact you can have on other people is huge. I've never had the opportunity to do something for my community and EmpowHER gave me that"
>
> – Young person

A total impact of 1.3 over 4.4 months may seem high, when compared to the 0.5 life satisfaction change when moving from unemployment into employment, which is considered one of the most important life changes for wellbeing. However, an impact of 1.3 life satisfaction points is in line with what could be expected, given what is known from other, broadly similar, programmes. For example, Action for Happiness runs a six-week course (for two hours each week) which focuses on understanding some of the important building blocks for happiness and developing pro-social behaviours. The Randomised Control Trial for this course found an impact of 1.0 life satisfaction points over the period in question among a general population (with impact sustained and even increasing after two months).[101] Given that EmpowHER targets a population with lower wellbeing,[102] one would expect similar or even stronger impacts. Indeed, the EmpowHER evaluation showed that those with lower starting wellbeing demonstrated on average the highest increases in wellbeing.[103]

Allowing for the observed completion rate among participants, we estimate the value of these wellbeing benefits to be around £6,538 per participant.[104]
This would represent a total monetised wellbeing benefit of £58,839 per group of nine YW&G.[105] Allowing for typical drop-out rates for participants, this represents a cost of around £11,500 per average group size of nine YW&G.
This means that the benefits of the programme outweigh the costs for each group.

(101) Krekel et al (2021): A local community course that raises wellbeing and pro-sociality: Evidence from a randomised controlled trial

(102) For EmpowHER, the average starting wellbeing was 1.1 life satisfaction points below UK average for girls in cohort 1 and1.5 life satisfaction points lower in cohort 2. UK average for girls comes from The Children's Society Annual Household Survey.

(103) It is a reasonable assumption that the benefits of EmpowHER could, as with Action for Happiness, continue for an additional two months after participation ends. This is explored in the sensitivity analysis.

(104) Compared to a benefit of £7,562 per completed participant in 2024 prices.

(105) Recommended group size was seven to 12, payment was provided based on a group size of eight to 10, with flexibility from the delivery partners in how this would be split. Groups were usually organised in similar age categories. Calculation: total programme costs split by all participants leads to a cost per participant of £1,520 in 2024 prices.

Overall, our analysis highlights that EmpowHER is likely to deliver good value for money – providing around £5 of wellbeing benefits per £1 spent. This broad conclusion remains valid across a range of alternative scenarios, summarised in the technical annex to this section of the report.

# Conclusion

**Our analysis provides evidence that the EmpowHER programme is likely to have delivered good value for money. This further strengthens the case for investing in social action projects as a way of supporting young people at risk of low wellbeing.**

The results also demonstrate the value of having matched data – and the additional understanding about a programme when it is possible to estimate what could otherwise have happened, in the absence of the intervention.

This case study is a good example of the value of assessment and learning. UK Youth was interested to measure wellbeing and test how social action could have an impact on the wellbeing of YW&G with lower wellbeing. They intentionally paused after each cohort to learn how the EmpowHER programme was achieving its aims. The information on participation and wellbeing benefits were able to help the charity to steer their intervention for future cohorts.

# EmpowHER technical annexes

# Annex B1 – Sensitivity analysis

## Sensitivity test 1: what if the benefits lasted for a larger or smaller number of months?

The central case assumes that wellbeing benefits lasted for 4.4 months,[106] based on the average length between pre and post surveys for those who completed the minimum of three months. It could be assumed that the benefits lasted for only three months on average, since this was the minimum time period of engagement in EmpowHER to be considered a 'completed participant'. Or, it could be assumed that the benefits lasted on average for six months for the YW&G who completed the programme and engaged for more than 12 weeks.

If the benefits only lasted for 12 weeks, the minimum time period of engagement in EmpowHER to be considered a 'completed participant', the programme would create £3.50 worth of wellbeing benefit for every £1 spent.

If the benefits lasted for six months, EmpowHER would create almost £7 worth of wellbeing benefit for every £1 of spend. Wellbeing benefits lasting for six months is highly plausible, given that i.) EmpowHER aims to target the underlying causes of low wellbeing and ii.) High quality, causal evidence from a similar course, Action for Happiness[107], shows that the wellbeing benefit was maintained or even increased two months after the course ends.

(106) Specifically, we assume that wellbeing increases linearly from the pre-level to the level measured at the 4.4-month point, then decreases linearly from the 4.4-month point down to the wellbeing level experienced in the counterfactual group.

(107) Krekel et al (2021)

## Sensitivity test 2: what is the change in wellbeing needed per participant for EmpowHER to 'break even' in monetised wellbeing terms?

EmpowHER would only have had to led to an improvement in wellbeing for three months of 0.4 compared to baseline for the programme to break even (0.1 if the impacts lasted for a whole year or 0.2 if the impact only lasted six months). This seems likely, given the high estimated average increase in wellbeing from taking part in the programme and the modelled drop in wellbeing which could have taken place in the counterfactual.

## Sensitivity test 3: what if there was no drop in wellbeing for those who didn't participate?

If the wellbeing of participants remained stable without participation in EmpowHER, they would have benefited on average from a 0.9 increase in wellbeing. If the benefit lasted for 4.4 months on average, wellbeing benefits would be estimated at £3.70 per £1 of spend.

If wellbeing benefits lasted for six months, every £1 invested in EmpowHER could have resulted in just over £5 of wellbeing benefit.

## Sensitivity test 4: what if the costs were higher or lower?

For example, only the costs associated with delivery could be included, excluding the costs associated with developing the programme, plus excluding a proportion of the costs associated with write-up and learning. This could be an estimate of the cost of 'further roll-out', since the programme design has already largely been completed. In this case, the cost per completed participant could drop as low as £1,240. Since the costs associated with achieving the same wellbeing benefit are lower, this would increase the ratio of the benefits per £1 of cost to just over £6, compared to £5 in our central scenario.

On the other hand, full resource costs may have been higher in the EmpowHER model delivered between 2018 and 2021 if there were significant resources provided free of cost. As an example, these could have represented an additional 5% of costs, which would reduce the benefit cost ratio to £4.90 of benefits per £1 of spend (from £5.10 in the central scenario).

In a higher cost scenario, the costs of all cohorts could be included, even though no wellbeing benefit of cohort 3 is included, the cohort impacted by Covid. This is a very high cost scenario, an example of a (fairly extreme) challenge which programmes could come across. In this unlikely situation, costs per completed participant would be £1,740, resulting in a benefit of just over £4 for every £1 spent.

# Annex B2 - Data which was required for wellbeing value for money analysis

- Life satisfaction before participation with date of interview.
- Life satisfaction on completing participation with date of interview.
- Demographic information, including on barriers.
- Participation information.

**Matched**

- Total number of participants and those who completed programme.
- Costs, including match funding, split by calendar or financial year. Split of costs by cohort and phase (delivery or learning).
- Data from Understanding Society Youth Panel (available from UK Data Service).

# Annex C: Get Out Get Active

## Introduction and overview

Get Out Get Active (GOGA), created and led by Activity Alliance,[108] is a programme that supports disabled and non-disabled people to enjoy being active together. GOGA aims to get some of the UK's least active people moving more through fun and inclusive activities.

Activity Alliance is the national charity and leading voice for disabled people in sport and activities. Their mission is to improve opportunities to be active, empowering disabled people to get involved in sport and activities in the way they choose. They work with disabled people and an alliance of committed partners to do this and build a movement for change.[109]

GOGA received a total of £9.5 million funding from Spirit of 2012, London Marathon Foundation, and Sport England between 2016 and 2024, with over 35,000 regular participants during this time. Over a third of participants were disabled.

This report assesses the wellbeing value for money of GOGA between 2016 and 2024. Since matched data was gathered for a sample of participants, it was possible to create a matched group to again assess how wellbeing may have changed in the absence of participation.

(108) Funded by Spirit of 2012, with additional investment from Sport England and London Marathon Foundation.

(109) Established in September 1998, Activity Alliance's vision is fairness for disabled people in sport and activity. Activity Alliance is the operating name for the English Federation of Disability Sport (EFDS).

## We found that:

- Disabled participants of GOGA were estimated to experience an average 0.05-point improvement in life satisfaction[110] after an average of 6.6 months (see Table 4).
  Since the wellbeing of disabled non-participants could have decreased by around 0.38 in this same period, this represents a total of 0.43 life satisfaction points over 6.6 months.
- Comparing the total resource costs to the wellbeing benefits of this group gives an estimate that the programme could have resulted in £3.70 of wellbeing benefits per £1 of cost.
- There are uncertainties in the scale of this benefit, given the assumptions in each step of the analysis. However, approaches such as weighting the data and the use of a matched comparison group give us reasonable confidence in the findings.
- These findings are for a timeframe of six months, which is the time over which sufficient wellbeing data was gathered. However, since the GOGA programme focused on getting and keeping people active, as well as training staff and organisations to be truly inclusive, there are likely to be longer-term benefits to wellbeing, including longevity, which do not arise within this timeframe.[111]

(110) We weighted the life satisfaction estimate to reflect demographics and characteristics of all participants, compared to sample. For this reason, the figure is different from pre-post change in life satisfaction reported in the GOGA1 and GOGA2 evaluation reports. See Tables 3 and below

(111) The full impact on wellbeing includes changes to quality of life (levels of wellbeing) as well as length of life.

**This case study, again, shows the value of creating a matched group, to estimate how wellbeing may have otherwise changed. This gives us more confidence in the scale of the wellbeing benefit which can be attributed to the programme and ultimately a benefit cost ratio: a comparison of the monetised wellbeing benefits to the monetised financial costs.**

It also demonstrates that sampling – gathering data from some, but not all, of the participants in a programme – can be a pragmatic approach for gathering wellbeing evidence, but there should be caution and sufficient consideration to ensure the sample is representative.

Lastly, it highlights the importance of longer-term follow-up, to understand the full wellbeing impacts of a programme. This is especially relevant when the aims of a programme are structural changes or lifetime impacts. For example, with specific evidence of how an inclusive approach impacts on wellbeing,[112] and longer-term evidence of how the GOGA approach changed inclusivity in delivery, it could be possible to extrapolate and estimate potential longer-term wellbeing impacts. In addition, evidence of changes to physical activity five or 10 years later could be combined with evidence of reduced mortality risk to estimate changes in length of life, with associated wellbeing benefits. This was not yet possible to incorporate in the wellbeing value for money case study. We have lower confidence in these types of extrapolations of what could happen compared to evaluations of what did happen, but it can provide a fuller picture in sensitivity analysis.

(112) From specific evidence designed to draw out the contribution of inclusivity, rather than physical activity, connection, or other drivers.

## Measuring wellbeing lesson 5: interpreting results

- In some cases, maintaining wellbeing can be a desired wellbeing benefit. This case study shows there was no significant improvement in pre-post wellbeing for the group in focus (disabled participants), but there was a significant improvement compared to what would have otherwise happened. The wellbeing of the matched comparison group dropped significantly in this period, suggesting that the programme had important protective characteristics. This highlights the importance of correctly interpreting the results, understanding change (or a lack of change) in the broader context.

## Measuring wellbeing lesson 6: sampling

- Sampling is an effective approach to gather important information in a less burdensome way. Where it is not possible for the sample to be representative of a full body of participants, weighting is possible. Demographic and other important information needs to be collected for the sample and full group of participants.

## Measuring wellbeing lesson 7: longer-term follow-up

- Some programmes or projects are seeking to achieve structural changes and lifetime impacts on wellbeing. To estimate the full impact on wellbeing, it may require a specific consideration of the pathways and the evidence linking these pathways to wellbeing, with longer-term follow-up.

# Background

**One in four of us do fewer than 30 minutes of physical activity a week, while one in six deaths is caused by inactivity.[113] The costs to the state of physical inactivity are high. According to the Everybody Active, Every Day governmental report, inactivity costs the UK an estimated £7.4 billion each year.[114]**

Disabled people are half as likely to be active as non-disabled people[115] and report lower levels of wellbeing than non-disabled people. The ONS reports average life satisfaction estimates of 6.7 compared to 7.9 for non-disabled people.[116] A higher proportion of disabled people feel lonely compared with non-disabled people.

Activity Alliance's research found that two-thirds of disabled people want to take part in activity with non-disabled peers[117] and the Annual Disability and Activity Survey in 2022-23[118] found "both disabled and non-disabled people want to take part in group activities with a mix of disabled and non-disabled people or have no preference (78% and 90%)", showing support for inclusive activities.

**1/4** One in four of us do fewer than 30 minutes of physical activity a week.

**1/6** One in six deaths is caused by inactivity.

**£7.4bn** Inactivity costs the UK an estimated £7.4 billion each year.

(113) Sport England: Research and Data: Inactive people. Accessed March 5, 2025

(114) Public Health England (2017): Everybody Active, Every Day: Two years on. Accessed March 5, 2025

(115) Data from Sport England's ongoing Active Lives survey (2021-22) highlights how disabled people or those with long-term health conditions are twice as likely to be physically inactive, with 41% inactive compared to 21% of non-disabled people. This activity gap has persisted over the years, and only slightly reduced from 43% and 21% respectively in the first year of data collection (2015-16).

(116) ONS (2019) Disability, well-being and loneliness, UK: 2019. Accessed February 18, 2025

(117) Activity Alliance (2019) Taking part with disabled people. Accessed March 5, 2025

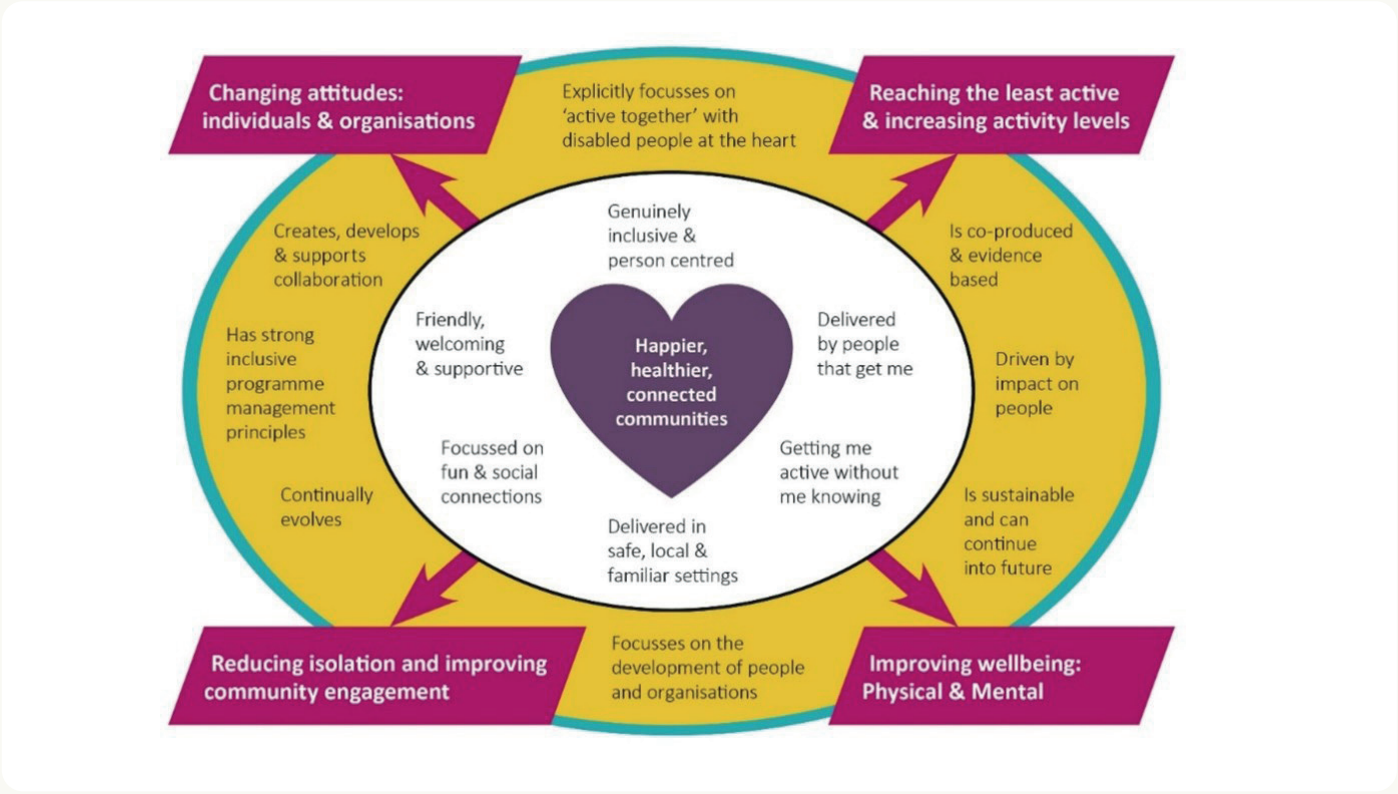
(118) Activity Alliance (2023) Annual Disability and Activity Survey. Accessed March 5, 2025

# The Get Out Get Active programme

**GOGA is a programme created to bring disabled and non-disabled people together to be active. It looks to engage the least active communities in fun and inclusive ways. Taking place in 21 locations across the UK, GOGA's lead partner, Activity Alliance, teamed up with a range of partners to help reach more people through local and national expertise.**

GOGA is about much more than being active. It strengthens community spirit, increases confidence, and improves mental health. The success of GOGA comes from tapping into people's real life motivations to be physically active, underpinned by Activity Alliance's research and principles.[119] The GOGA approach and a simplified representation of the GOGA steps to impact are shown in Figures 9 and 10.

Figure 9: The GOGA approach




(119) Activity Alliance (2014) Talk to me. Accessed March 5, 2025

Figure 10: GOGA simplified logic model

| Inputs | Activities | Outputs | Outcomes | Impact |
|---|---|---|---|---|
| Funding | Community consultation and co-production | Over 35,000 participants (of whom over a third were disabled and 1 in 4physically inactive when joining GOGA) | Increases in physical activity | Improved wellbeing |
| Match funding | Delivering genuinely inclusive existing or new activities | 4,118 volunteers | Reduction in inactivity levels(120) | |
| Expert local and national expertise to shape delivery | Outreach/ engagement of participants and volunteers | Staff and volunteers trained | Improved social connectedness(121) | |
| | Training of workforce | | Improved perceptions of disabled people(122) | |

(120) Wavehill (2024): Evaluation of Get out Get Active Phase 2 Final report. Accessed February 18, 2025, reports that GOGA has supported 78% of those physically active to do more physical activity, and 74% have continued to be active and maintained their activity levels outside of the GOGA programme.

(121) Wavehill (2024) reports that six in 10 are now more connected in their community.

(122) Wavehill (2024) reports that 65% say they have a more positive view of disabled people, with 58% of interviewees at nine and 15 months after taking part in the GOGA programme say that their view of disabled people is more positive as a result of their programme participation.

# Our approach

We linked the wellbeing data (life satisfaction) through to the wellbeing value for money via the following six steps:

## Step 1: estimate the average change in life satisfaction in those who participated in GOGA

As described above, improving wellbeing was one of the aims of GOGA. Spirit of 2012 required that all funded projects collected consistent data on wellbeing, including life satisfaction. Wavehill evaluated GOGA. They collected data on participants' life satisfaction at baseline (up to a month after starting the programme) and just over six months later. This was collected for a sample of participants who agreed to answer the fuller questions. This data was used to estimate the average change in life satisfaction.

The sample self-selected and a comparison of the demographic data shows that they had different characteristics than those of the full population of participants[123] (see Table 2). Table 3 demonstrates that certain groups who are under- or over-represented in the sample reported a higher or lower change in wellbeing, which would lead to a biased estimate of the change for all participants. This means that the change in life satisfaction reported by the sample would give a biased estimate compared to the change in wellbeing which would have been reported by the full population of participants. To address this, estimates were weighted in the following step.

(123) As estimated using information from Upshot, the initial starting survey which was filled out by all participants in the majority of local authorities. Data from Upshot was available for ~75% of total participants.

Table 2: Demographic and other characteristics: comparison of sample and full population

| | GOGA1 | | GOGA2 | |
|---|---|---|---|---|
| | Sample | All participants | Sample | All participants |
| Gender: Female | 73% | 61% | 66% | 59% |
| Gender: Transgender, non-binary, other, prefer not to say, not provided | 0% | 0.4% | 0.4% | 3% |
| Disability | 56% | 40% | 55% | 35% |
| Age 15-30 years | 6% | 22% | 8% | 17% |
| Age 31-50 years | 21% | 30% | 15% | 22% |
| Age 51-64 years | 22% | 17% | 30% | 17% |
| Age 65 + | 23% | 15% | 43% | 21% |
| White | 86% | 76% | 84% | 79% |
| Asian/Asian British | 6% | 13% | 7% | 9% |
| Still engaged in project (Y) | 64% | Data not available, assume same ratio between all participants and sample as for GOGA2 32% | 39% | 20% |

Table 3: Sample reported average life satisfaction change, split by different groups

| | Average change in LS – GOGA1 | Average change in LS – GOGA2 |
|---|---|---|
| All | 0.25 | 0.17 |
| | | |
| Female | 0.20 | 0.18 |
| Male | 0.58 | 0.15 |
| | | |
| Disabled | 0.26 | **0.44** |
| Not disabled | 0.36 | -0.13 |
| | | |
| 15-30 years | -0.2 | 0 |
| 31-50 years | 0.44 | 0.23 |
| 51-64 years | 0.09 | 0.01 |
| 65+ years | 0.51 | 0.29 |
| | | |
| Not engaged at second interview | -0.02 | 0.32 |
| Still engaged at second interview | 0.49 | -0.06 |

We examine the changes in wellbeing for all participants and also for disabled participants, weighted and unweighted.

## Step 2: estimate what change in wellbeing may have taken place without GOGA

Here we seek to understand how wellbeing may have changed for participants if they didn't take part in GOGA, ultimately to understand the full impact of GOGA, in the following steps. We estimated the change in wellbeing of a matched comparison group using Understanding Society, a panel survey.

We used the following information on participants in order to find a matched group in the Understanding Society panel:

- Sex.
- Age.
- Ethnicity.
- Starting wellbeing score.
- Calendar year.

A matched comparison group was created of all participants, of disabled participants, and also weighted.

## Step 3: estimate the change in wellbeing due to GOGA

By comparing the difference in wellbeing between the matched group and the group who participated in GOGA, we can estimate the wellbeing impact of participation in the programme.

### Table 4: Results from matching analysis for GOGA (124)

| Group | n | Matching method | Change treatment | Change control | Impact on LS | p-val |
|---|---|---|---|---|---|---|
| All | 373 | Propensity score | 0.20 | 0.08 | 0.12 | 0.38 |
| All | 373 | Mahalanobis | 0.20 | 0.00 | 0.20 | 0.12 |
| All weighted | 373 | Mahalanobis | 0.02 | -0.09 | 0.08 | 0.68 |
| Disabled | 217 | Propensity score | 0.31 | -0.12 | 0.44 | 0.03 |
| Disabled | 217 | Mahalanobis | 0.31 | -0.12 | 0.44 | 0.02 |
| Disabled weighted | 217 | Mahalanobis | 0.05 | -0.38 | 0.43 | 0.04 |

We focused where the results were statistically significant, that is to say, where it is likely that the difference is not due to chance but rather due to differences in the two groups. This means we focused on disabled participants, for whom wellbeing increased slightly over the 6.6 months, while wellbeing of a group with similar characteristics was estimated to have dropped in the same time period. Since GOGA was delivered during some of the years where 'usual' social activities were disrupted by Covid and continued to be delivered in online formats, it suggests that GOGA may have had protective characteristics, halting the drop in wellbeing which may have been observed for others over this time.

We assessed the impact over 6.6 months, the average time between the pre and post surveys.

To calculate the number of WELLBYs, which are a 1-point change in life satisfaction (0-10 scale) per person per year, we need to apportion the figure by the number of months.

(124) Consistent with the other case study, we assume a non-linear translation from the 7-point scale to the 11-point adult ONS life satisfaction scale that is based on aligning the cumulative distributions for all adults in 2021 between Annual Population Survey (1-point life satisfaction measure) and Understanding Society (7-point life satisfaction measure).

Calculation:

(Change in treatment group - change in control group) * length of days/days in a year

= (0.05 - -0.38) * 198/365
= 0.43 * 6.6/12
= 0.23 WELLBYs

We multiplied this wellbeing benefit by all 13,506 disabled participants over the eight years.

## Step 4: monetise this change in wellbeing

The change in wellbeing is multiplied by the monetised value of a 1-point change in life satisfaction, commonly referred to as a WELLBY.

The value of a WELLBY comes from the HMT Green Book Supplementary Guidance on Wellbeing. These values are uprated to 2024 prices, meaning that a single WELLBY has a central estimate of £15,900 (in 2024 prices).[125]

This gives total wellbeing benefits (monetised in 2024 prices) of £50.1 million.

## Step 5: estimate the costs of the programme

GOGA reported spend over the eight years, including staff costs, overheads, further resources, and wider costs such as evaluation. This included Spirit of 2012 funding as well as matched funding. Additional resources provided 'in kind', such as venues, catering, or staff training provided for free, were given their 'equivalent' cost and included in the total cost estimates.[126] This includes the costs for all participants, not just disabled participants.

We expressed all costs in 2024 prices.

This gives a total programme cost of £13.4 million in 2024 prices (including full resource costs).

It should be clear that this is not the funding received by Spirit of 2012 and match funders,[127] nor the total amount which was spent on the project over the eight years. Rather, it is:

(125) Rounded to nearest £100.

(126) For value for money estimates, we aim to compare the total resource costs to the benefits. Financial costs are a subset of the total resource costs.

(127) Which was £9.5 million

- The total value of the project spend including the value of resources received in kind. This gives us the total **resource cost**, which is the full cost of the programme and the appropriate figure to use for value for money analysis.
- What the value of this spend would be if it were spent 'today'. Due to inflation, £100 today can buy us less than we could have bought for £100 in 2016. This means that £100 spent in 2016 has more value (it was able to buy more) than £100 today. We put all costs into the same 'price year' so we can compare them equally, i.e. £1 has the same value no matter when it was spent.

## Step 6: assess the value for money

The costs of delivery of the programme were compared to the wellbeing benefits monetised in Step 4. This gives an estimate of the benefits produced for every £1 spent: an assessment of value for money. We applied sensitivity analysis around this figure.

## Key assumptions

- We assume that the data gathered for the 217 disabled participants for whom we had full data to enable a reasonable match was representative of all 13,506 disabled participants in GOGA. Although the evaluators made initial attempts at ensuring a balanced distribution in terms of the different localities, they ended up interviewing everyone who indicated they were open to be surveyed, to ensure they had a sufficient sample. Those willing to be surveyed may not have the same characteristics of the wider participant pool. This is confirmed through comparing the sample demographics to those of all participants, information available from the initial starting survey which was filled out by all.[128] The sample has a different age profile, a higher proportion of females, a higher proportion of those reporting white ethnicity, and a greater proportion of those with a disability (see Table 2). In addition, the sample had a higher proportion who were still engaged at the six-month point compared to the full population of participants.[129]

- We increased our confidence in these estimates by weighting the estimates to align the demographic characteristics such as sex, age, and ethnicity, as well as whether they were still engaged at the six-month follow-up interview. However, there may have been further, unmeasured characteristics which influenced self-selection into the wellbeing questionnaire as well as the reported change in wellbeing.

- The average length of time between the pre and post survey was 6.6 months. We assume that, on average, wellbeing falls again after this point. However, there could have been changes due to GOGA which led to a longer lasting impact on the life satisfaction of participants. GOGA was designed to enable people to progress, to leave their GOGA group but to continue engagement and fun through sport.

(128) Through sign-up to the 'Upshot' data management system, where participants could indicate if they were open to be surveyed or not.

(129) For GOGA2, 38% of the sample were engaged at six months, compared to 20% of the full participant population. For GOGA1, 58% of the sample were still engaged at six months, but we cannot compare this to the full participant population due to the legal requirement for data to be deleted after a number of years.

- We only have the data at year intervals for the matched control group, i.e. the change in wellbeing which took place over 12 months for a similar group. However, to compare the two groups, we need to estimate what change may have happened at 6.6 months for our central scenario. For this, we assume that the change over one year for the comparison group is linear and that the change at month 6.6 is equivalent to 6.6/12ths of the change observed for the whole year.

- There could have been changes outside of the programme which influenced the wellbeing of participants. Even though we have created a matched sample, which aims to 'control' for these outside influences which could have impacted the participants and non-participants alike, there may be factors relating to the programme. These caveats add some uncertainty, and we are not currently able to minimise it further given the evidence available. However, we explore the robustness of our core results to variations in these assumptions as part of our sensitivity analysis.

- Lastly, we believe the estimates may be conservative, for a few reasons. Firstly, there could be potential cost savings to the health systems from improved outcomes, such as reduced mental or physical health costs, that are not captured in our analysis due to insufficient evidence. In addition, we are missing any longer-term wellbeing impacts due to physical activity's impact on mortality risk. The full change in wellbeing from a programme depends on i.) the quality and ii.) length of life, i.e. i.) the change in reported wellbeing and ii.) how a programme may influence mortality. This second aspect was not possible to include in our calculations, but could be particularly relevant in this case study, given the evidence that physical activity reduces mortality risk[130] and would, in effect, increase years of life.

(130) See for example Kelly et al. (2014): Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship

# Results

## Benefits

**Disabled participants of GOGA experience an average 0.05-point improvement in life satisfaction after an average of 6.6 months using data gathered between 2017 and 2022. Since the wellbeing of non-participants could have decreased by around 0.38 in this same period,[131] this represents a total of 0.43 life satisfaction points over 6.6 months.**

Assuming that the benefits last for 6.6 months, the average improvement in wellbeing life years due to GOGA could be around 0.23 per participant. This has a value of around £3,700 per disabled participant when monetised using the UK government's guidance. This suggests a total wellbeing benefit of £50.1 million for the 13,500 disabled participants in GOGA over the lifetime of the programme.

The total costs of GOGA over the eight years of delivery was £13.4 million in 2024 prices.[132] This includes funding from Spirit of 2012, match funding, and an estimate of the 'in kind' resources such as venues provided for free, staff training, and expertise provided for free. This is an average cost of £375 per participant.[133]

(131) Analysis was carried out for the change over a year, an estimated 0.8 drop over the year.

(132) £9.5 million in nominal terms.

(133) This cost cannot be directly compared to the estimated benefits per disabled person as not all participants were disabled. When these total costs are split only by disabled participants (since we are only including the wellbeing benefits of disabled participants), the cost per disabled participant is £993. This is not to say that it was more expensive to deliver GOGA to disabled participants, rather the split of the costs among a smaller or larger group of participants.

Comparing the total resource costs of £13.4 million to total wellbeing benefits gives us an estimate that the programme could have resulted in £3.70 of wellbeing benefits per £1 of cost.

Given the assumptions above in each of the steps of the analysis, we cannot be certain of the scale of this benefit. However, with the weighted analysis and the use of the matched group, there is much higher confidence than we otherwise would have. That wellbeing dropped for a created comparison group highlights the potential protective characteristics of GOGA and shows the benefits of trying to understand what otherwise could have happened. In addition, we have explored a range of alternative assumptions in our sensitivity analysis (summarised in the technical annex). These demonstrate that it remains plausible that the benefits of the GOGA programme outweighed its costs in a wide range of different scenarios.

# Conclusion

## Summary of key findings

**Our analysis has highlighted that GOGA is likely to have had a 'protective' wellbeing impact for disabled participants. This impact is likely to mean that the benefits of the programme outweigh its costs.**

We cannot have full confidence in this assessment due to the data availability and necessary assumptions. However, sensitivity testing estimates that only around 40% of disabled participants would have had to have received these wellbeing benefits for the monetised benefits to be greater than the costs. Including wider potential benefits such as volunteering and lengthening the assumption on the duration of benefits strengthens this value for money case further. In addition, we currently include no wellbeing benefits for non-disabled participants.[134] However, it could be the case that certain sub-groups of non-disabled participants may have also benefited from a statistically significant improvement in wellbeing which was not possible to discern at the aggregate level. Lastly, since the 'pre' wellbeing questions were asked up to one month after starting participation in GOGA, they could underestimate any improvement in wellbeing, if there was an initial improvement in wellbeing during the first few weeks of participation. If this was the case, the analysis would be an underestimate of the wellbeing value for money.

This case study, again, demonstrates the further confidence we can have in the change – attributed to the programme – where we have sufficient matched data and can create a matched comparison group. This allows us to monetise the wellbeing benefits and express the value of these compared to the costs in a benefit cost ratio. It also highlights that grantees don't need to ask every single participant for outcome data. Sampling can be an appropriate method for gathering wellbeing data, where we are confident that the sample is representative. Where this is not the case, further steps such as weighting are necessary.

(134) Given that the weighted analysis found no statistically significant change for this group, on average.

# GOGA technical annexes

# Annex C1 – Sensitivity analysis

## Sensitivity test 1: what happens if the benefits lasted for a larger number of months?

In our central case, we assumed that wellbeing benefits last for 6.6 months,[135] based on the average length between pre and post surveys. We could assume that the benefits lasted on average for, for example, nine months. If this were the case, GOGA would create around £5.20 worth of wellbeing benefit for every £1 of spend.

## Sensitivity test 2: what is the change in wellbeing needed per participant for GOGA to 'break even' in monetised wellbeing terms?

With costs of around £375 per participant, GOGA would only have needed to have improved wellbeing by 0.04 on average **for each of the 35,771 participants**, over 6.6 months, to break even. With the focus on disabled participants, GOGA would only have needed to have improved wellbeing by 0.1 compared to the counterfactual to break even.[136]

(135) Specifically, we assume that wellbeing increases linearly from the pre-level to the level measured at the 6.6-month point, then decreases linearly from the 6.6-month point down to the wellbeing level experienced in the counterfactual group.

(136) Again, over 6.6 months, the average length between pre and post surveys which is used in the central scenario. This represents 0.02 WELLBYs for all participants and 0.06 WELLBYs for only disabled participants.

## Sensitivity test 3: what proportion of disabled participants would have had to have received this benefit for the programme to break even?

In our central case, we apply the estimated wellbeing benefits to all 13,506 disabled participants. However, even with the weighting which was carried out, there are remaining uncertainties of whether the findings from the sampled participants are representative of the wider pool of disabled participants. To test the robustness, the analysis tested the proportion of participants who would have needed to receive these benefits for the programme to break even. Only 27% of disabled participants would have had to receive the estimated benefit of 0.43 life satisfaction benefit over the 6.6 months for the project to break even.

## Sensitivity test 4: what if there were additional wellbeing benefits from volunteering?

GOGA did not separately estimate the wellbeing benefits gained by volunteers in the programme. However, other recent studies have shown that volunteering is beneficial for wellbeing, with a causal impact of around 0.2 life satisfaction points.[137] Based on this and the figure of 4,118 volunteers over the eight years, we can estimate the potential additional wellbeing benefit from volunteering as part of GOGA. The 28% of volunteers who were disabled were excluded, to avoid double counting.[138] We multiplied the remaining figure by the 72% who were classified as regular volunteers who volunteered for at least six months, giving a total of 2,135 regular volunteers who were non-disabled.[139] Including this rough estimate increases the benefit slightly from £3.70 of benefits per £1 to £4 of wellbeing benefits.

(137) For biweekly volunteering as part of an NHS responders scheme. See Krekel et al (2024): Happy to Help: Welfare Effects of a Nationwide Volunteering Programme.

(138) Since we include the wellbeing of disabled participants in our estimates and volunteers are counted as participants.

(139) We multiply this by 0.2 over an estimated timeframe of six months to give a total WELLBY estimate of 312.

## Sensitivity test 5: what if there were fewer disabled participants?

To arrive at the figure of 13,506 disabled participants, we used the figure of 35,711 GOGA participants from the final impact reporting[140] and multiplied by 38%[141], the proportion of disabled participants from Upshot registration data (n=26,990 i.e. 75% of total participants). The total participant number includes data from Upshot, the registration system used across the UK by the evaluators, as well as data from local authorities who had not signed up to Upshot. We only had information on disability from the Upshot data. For the further participants, we estimated the proportion of disabled participants from the Upshot data. This means it could be an under- or over- estimate. We could take a more conservative estimate and only include the number of disabled participants who registered through Upshot. This would give us a figure of 10,095 disabled participants, which would reduce the wellbeing benefit to £2.80 per £1 invested in GOGA.

# Annex C2 – Data required for wellbeing value for money analysis

- For the sample: matched pre and post reported life satisfaction with date of interviews and demographic information.
- Data from Understanding Society Youth panel (available from UK Data Service).
- For all participants: number of completed participants with information on level of involvement and demographics.
- Costs, including match funding, split by calendar or financial year. Split of costs by cohort and phase (delivery or learning).

(140) Wavehill and Activity Alliance (2024) Get Out Get Active Impact Report. Accessed November 5, 2024

(141) 40% for GOGA1 and 35% for GOGA2

# Checklist and important links

## Checklist for wellbeing value for money analysis

This is a checklist if you want to assess the wellbeing value for money of investments. There are wider points to consider for evaluation in general and wellbeing measurement in general.

Going further down the checklist will enable a higher level of confidence in quantitative value for money estimates. However, the full checklist will not be relevant for every organisation nor every project. Even if you are not sure if you will be conducting a value for money analysis straightaway, these steps will mean that you or an external evaluator can use the data in the future to make a confident assessment.

Each of these steps is useful in its own right.

### Getting started

- Establish a theory of change, of how the proposed activities are expected to lead to a change in wellbeing for different groups.[142] Draw from existing evidence to understand where there may be gaps.
- Think through what is important to understand, and plan this into your evaluation from the beginning. This will influence the information you collect and in what format.
- Plan how you will record costs. Value for money estimates rest upon not only the benefits, but an accurate record of costs. Record all the costs that go into running your project, including set-up and running costs, overheads, as well as support from others. When it comes to funding your project, add up all the funds that will pay for your project, from all sources. Think about the headings you give these costs in a spreadsheet – are they simple to understand for someone not currently involved in the project?
- If you are commissioning an external evaluator, explain that you want wellbeing value for money analysis to be part of your project, and ask them how they will go about this.

(142) And the intermediate steps, of outputs and outcomes.

## Baselining

This enables organisations to understand their beneficiaries, with the opportunity to amend their approach on the back of this (see EmpowHER case study). Where relevant for the programme, it can also provide the option for individual participants to reflect on any change and their journey.

- Ask the ONS life satisfaction question to participants at the beginning of your programme, ideally as a suite of further questions to understand their wider context and drivers of wellbeing. It is also important to record wider demographic questions such as age, ethnicity, sex, whether the participant has a disability.[143] Record the date of the survey.
- Detail: wellbeing questions should be asked after core questions on household and individual demographics, but before any wider questions on health, employment, lifestyle, or programme participation. This is based on findings from ONS cognitive testing. It allows time for rapport to be built up between the interviewer and the respondent without allowing later questions to influence response to the personal wellbeing questions.
- To give the future possibility to 'go further' and understand change, make sure each survey response has a unique identifier so it can be matched to a future response.
- Compare to national benchmarking data to understand if you are reaching the 'right' people (i.e. for improving wellbeing, those with lower starting wellbeing; if you are targeting a particular driver of low wellbeing such as loneliness, those with higher levels of loneliness). Targeting the intended beneficiaries is an important value for money consideration in its own right.

(143) If possible: employment status, in a partnership or not, family income, or socioeconomic group.

## Understanding change

These steps are in addition to the steps in 'Baselining'. This enables organisations to understand the reported difference over the time period of a particular programme and estimate what change may have happened. For value for money analysis, it enables scenario analysis (see City to Sea case study).

- Record the number of participants, volunteers, and other groups who are anticipated to have a change in wellbeing.
- Monitor how each person has been involved, especially where programmes are delivered in different ways in different parts of the country and different levels of involvement are possible. This may include:
- Length of participation (e.g. number of weeks or months, or one-off).
- Frequency of participation.
- Volunteer hours and frequency.
- Was there a set programme and did they attend every week?
- This will depend on the programme.
- If your project involves sufficient engagement, for which you would expect more than a temporary change in wellbeing, ask the life satisfaction question (plus any others) when the programme ends.
- Detail: ask the question in the same format as at the beginning of the programme, i.e. if it was filled out individually on paper, do the same; if via an interview, do the same.
- Record the date of the post survey. Assign the same ID so it is clear which before and after scores go together.
- Use statistics expertise to understand the change, as part of the wider context.
- Is this change statistically significant?
- Representativeness really matters. If a sample is not representative, weighting can be done (see GOGA case study). It is important to consider:

- What proportion of people dropped out or didn't answer the post question?[144]
- If you used a sample, were they chosen to be representative (or rather self-selected)?
- What happened in the wider context? (Was wellbeing dropping or increasing in general in this time period? Was the pre result in winter and the post in summer, when life satisfaction scores are higher?)
- Have you carried out quality assurance? Errors and typos are common and when unchecked can influence the findings. Assumptions are often necessary, including whether findings from a sample can be applied to a wider group. Quality assurance should be used to check any statistical calculations, analysis, and assumptions.
- A statistically significant difference between pre and post, with this wider context, does not tell us the impact, i.e. whether the change is attributable to your programme, or whether wider factors were involved. There could have been wider changes taking place which influenced the difference in life satisfaction (see steps for 'Understanding impact'). However, it is an important step, for which we can carry out scenario analysis (see The Wave Project case study).
- Record your costs, alongside the year in which it they were spent.
- If you are using an external evaluator, use this guide to set some expectations for them, and ask them how they will present back to you on the above points.

(144) As a rough rule of thumb, we have less confidence when more than 40% of beneficiaries or 40% of the representatively chosen sample didn't answer the post question.

## Understanding impact

These steps and are in addition to the steps in 'Understanding change'. This enables organisations to create a matched comparison group and estimate the impact, i.e. greater confidence in the change which could be attributed to the programme. For value for money analysis, it enables a cost benefit ratio (see EmpowHER and GOGA case studies).

- Sufficient quality assurance of data collection.
- Match data pre and post alongside demographic characteristics.
- Minimum of data for 50 individuals (with both pre-post) and ideally more than 100.
- Representative reports: pre and post collected for at least 60% of participants to have full confidence. Where sampling, ensure an approach has been taken to be representative (and/or collect data to understand how representative this sample is of the beneficiaries as a whole).
- If you are using an external evaluator, use this guide to set some expectations for them, and ask them how they will present back to you on the above points.

### Estimating future changes

Some programmes may be planned to lead to longer-term changes in wellbeing, but in many cases it is not proportionate nor feasible to have a follow-up after five, 10, or even 20 years. The steps below can help organisations to express what some of these future changes may be. There will be much lower confidence in these estimates, since we are not measuring wellbeing directly. Such calculations will extrapolate what could happen, which is a lot more uncertain compared to evaluations of what did happen. However, it can provide useful scenarios of the potential full wellbeing implications.

- Develop the theory of change in further detail, to draw out the links to wellbeing in different timeframes for all the relevant groups. This may include changes in future physical activity, resulting changes in mortality risk, changes in likelihood of future employment, changes in likelihood of mental health conditions. Highlight the outcomes which are likely to lie outside of the timeframe of the evaluation.

- For the most important outcomes, develop an evaluation strategy: what needs to be measured within the timeframe of the evaluation to estimate future changes in wellbeing?

| | When you might want to do this stage | When you may decide *not* to do this |
|---|---|---|
| **Baselining – with wellbeing measures** | • To understand if the programme is targeting the intended beneficiaries, to give the opportunity to reflect on and amend approach. For example, do participants have lower wellbeing than 'average'? If you are targeting a particular driver of wellbeing such as loneliness, or physical inactivity, do participants report higher levels of loneliness/physical inactivity? If not, do you need to work with different organisations to include those who may benefit the most?<br>• To get a fuller picture of the participants' lives. By asking about their wellbeing alongside other relevant circumstances or aspects of their life you can check that your activities are still relevant. This could also pick up challenges for participants that you may not have considered in your planning.<br>• For some programmes, where there is a focus on participant reflection: to enable individual participants to reflect on their journey and any changes they have experienced. Recording this is helpful since we are bad at remembering – it can be difficult to remember how we previously felt. | • Where there are no follow-on programmes or cohorts and there is no potential for amending programming or approach.<br>• Where individual reflection is not planned in nor important for the programming. This type of reflection requires sufficient time, planning, and understanding.<br>• Where it is not important to have further information on the beneficiaries. (This may be the case for some shorter, one-off events, or where this information is already known e.g. participants have been referred from a programme which targets specific groups with low wellbeing.) |

| | | |
|---|---|---|
| **Understanding change in life satisfaction** | • To understand and estimate change over time, even when this change may not be possible to attribute to the programme or organisation.<br>• This guidance from the What Works Centre for Wellbeing provides more information.(145)<br>• To assess value for money through scenario analysis (see City to Sea case study).<br>• To understand distributional impacts, i.e. where those with lower starting wellbeing may have greater improvements in wellbeing, or particular groups may have differential impacts (requires matched data). | • Where the programme has not been designed to lead to meaningful changes in evaluative wellbeing (overall assessments of life).<br>• When there are too few participants to be able to get a statistically significant result. There is no definitive minimum level of data but typically we would encourage charities to aim for a minimum of 100 individuals with pre and post data in order to support economic evaluation.<br>• Where the reports are unlikely to be representative. This could be the case for, for example, when the percentage of responses to pre and post surveys is small compared to total beneficiaries (without a representative sampling approach). From our experience, a charity is doing well if it can gather data from at least 60% of its participants. |
| **Understanding life satisfaction impact** (using a matched comparison group) | • To estimate the change which was due to the programme.<br>• To carry out value for money analysis with benefit cost ratios (see GOGA and EmpowHER case studies). | • When there is no matched data for participants.<br>• When the number of participants is too small to give confidence in the impact, e.g. less than 50 with pre and post.(146)<br>• Where the reports are unlikely to be representative. |

(145) What Works Centre for Wellbeing: Analysing and interpreting your results – Evaluating wellbeing. Accessed March 20, 2025

(146) See for example, Green, S. B. (1991). How Many Subjects Does It Take To Do A Regression Analysis.

# Important and useful links

ONS Personal well-being user guidance

What Works Centre for Wellbeing Measuring your wellbeing impact: A practical guide for charities and social enterprises


Economics to improve lives

Jon Franklin
Info@pbe.co.uk

@pbe
@pbe.co.uk
www.pbe.co.uk
020 3632 2668