



Here for young people
Here for communities
Here for you

THE CENTRE FOR YOUTH IMPACT

PART OF YMCA ENGLAND & WALES

Standards of Evidence

For continuous quality improvement and
evaluation

Bethia McNeil and Charles Smith Ph.D, July 2025

Everyone should have a fair chance to discover who they are and what they can become.

About YMCA

YMCA believes in fairness and opportunity. There are essential building blocks for a full and rewarding life: a safe home; acceptance; guidance; friendship; physical and mental health; academic support; employment skills; and access to real opportunities. Many young people have never known these things; other people have lost one or more as they grew up, but we all need them. All of us. At YMCA, we provide these critical foundations for a fresh, strong start for young people and a better quality of life in the community.

This content was developed by the Centre for Youth Impact, part of YMCA England & Wales, with the support of a range of funders, including the Department for Culture, Media and Sport (DCMS), Paul Hamlyn Foundation and Esmée Fairbairn Foundation.

The contents and resources are made free and accessible under the terms of the Creative Commons [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) licence. Please review the terms before using or sharing.

The Creative Commons CC BY-NC-ND licence enables you to copy and distribute the material in any medium or format in unadapted form only, for non-commercial purposes only, and only so long as attribution is given to the Centre for Youth Impact at YMCA England & Wales.

BY: Credit must be given to the creator of the material

NC: Only non-commercial uses of the material are permitted

ND: No derivatives or adaptations of the work are permitted

Standards of Evidence

Introduction

These Standards of Evidence were developed by the Centre for Youth Impact, now part of YMCA England & Wales, through its work with research partner QTurn on ‘quality to outcomes’ design and method in evaluation. This model proposes a more nuanced, meaningful and practical approach to understanding and improving impact in informal and non-formal learning settings, grounded in young people’s experience of context and relational practice.

The Standards of Evidence were created to provide guidance to organisations (both delivery and infrastructure) and funders in the youth sector when they are preparing to invest resources in ‘performance’ measurement (that is, of quality *and* outcomes) for purposes of **continuous quality improvement** (CQI) or **evaluation**.

These Standards for Evidence are differentiated from other efforts to articulate hierarchies of evidence for the field, and professional researchers and evaluators, (such as What Works Centres or evidence clearinghouses) in that they are **designed with the needs and capacities of delivery organisations** as a starting point. Delivery organisations that undertake evaluations that comply with Standard 4 will be producing evidence that, when reported correctly, can be included in What Works-aligned frameworks under criteria for quasi-experimental designs.

Understanding continuous quality improvement and evaluation

CQI is an *internal* use of measurement and data for the purposes of improvement by practitioners. ‘Internal’ means that data and insight gathered is primarily for the practitioners and the organisation, rather than being undertaken for, or handed over to, an external agency such as a funder or commissioner. For CQI, data is most valuable as real time feedback to practitioners about their own practices and young people’s socio-emotional skills, without extensive aggregation and modelling. This data is mostly focused on **four causes of outcomes** in informal and non-formal learning provision for young people:

- ‘Dosage’ (the frequency and intensity with which young people engage with provision);

- Activity type;
- Provision fidelity/quality; and
- The level of young people's socio-emotional skills at baseline.

Standard 2 is particularly relevant to CQI.

In contrast to CQI, **evaluation** is focused on measurement of individual young people's skills at multiple time points. When the focus is limited to young people who participated in provision (i.e., who experienced the four 'causes' above), we refer to **outcome evaluation**: Data and results reflect the pattern of skill change (increase, maintain or decrease) for each individual young person at multiple timepoints, with reported differences for subgroups. Like data about causes, individual skill change data is valuable in disaggregated form (i.e., diagnostic information about individual young people) but these uses must be carefully considered due to confidentiality/anonymity promises required by ethics and GDPR. Outcomes evaluation results are also useful aggregated to the group level, particularly when pattern-centred methods are used to identify the groups.

Standard 3 is particularly relevant to outcomes evaluation.

When outcomes data for young people participating in provision is aggregated to the group level and compared to similar outcomes in other groups of young people, we refer to **impact evaluation**: Data and results reflecting the average pattern of skill change for one or more groups of young people who experienced the causes above is compared to another group(s) that did not experience the causes above (often referred to as a control or comparison group). Impact evaluation results can take a long time to produce (for example, because one has to wait until post-test when provision is almost completed) and are most valuable when aggregated and modelled according to specific rules of research design. Impact evaluation compares the average amount of individual change over time for specific groups, e.g., treatment or comparison, low or high fidelity/quality, low or high dose.

Standard 4 is particularly relevant to impact evaluation.

Glossary

Theory of change	<p>Description of anticipated patterns of causes and effects for changes in individual young people's socio-emotional skills.</p> <p>In informal and non-formal youth provision, the causes of young people's skill change are:</p> <ul style="list-style-type: none"> • dosage; • activity type; • provision fidelity/quality; and • young people's socio-emotional skill level at 'entry'. <p>The effects that follow are of two types: <i>outcomes for individuals</i> that reflect their change in skill and <i>impacts for groups</i> of individuals' average skill change.</p>
Fidelity/Quality	Description of what practitioners (adults who are either volunteers or staff members) need to do in order to create the causes of skill growth, assuring that the anticipated effects will follow.
Dosage	The 'amount' of provision a young person engages with/in, sometimes measured in hours, but also in relation to frequency and intensity (e.g. once per week for two hours, once a month for 30 minutes etc)
Outcomes	In these Standards of Evidence, an outcome is a change in an individual young person's socio-emotional skill, one type of effect caused by participating in high fidelity/quality youth provision.
Impact	In these Standards of Evidence, an impact is the difference between the average outcome for one or more groups who participated in high fidelity/quality provision and the average outcome for a different comparison group of young people who were exposed to something different, typically no provision or low-fidelity/quality provision
Alignment	The extent to which selected measures for provision (either of fidelity/quality or individual outcomes) are specifically focused on causes or effects named in the theory of change, i.e., well aligned (also known

	as 'valid')
Sensitivity	The extent to which the selected measures (and processes for using them) for individual outcomes are likely to capture real effects that actually occur for young people
Continuous quality improvement (CQI)	Focused on measurement of causes (see theory of change, above) using participatory methods and emphasising disaggregated data reported immediately after data collection, for the purposes of response, action and provision improvement
Outcomes evaluation	Data and results reflecting the pattern of skill change (increase, maintain, decrease) for each individual at multiple timepoints, with reported differences for subgroups
Impact evaluation	Data and results reflecting the <i>average</i> pattern of skill change for one or more groups of young people who experienced causes (see theory of change, above) compared to another group(s) that did not experience these causes.
CQI cycle	A structured and planned cycle of quality assessment, data review, improvement planning and action, followed by re-assessment. Such cycles often take place annually, and are undertaken internally by small teams of staff and/or volunteers.
Effect size	A quantitative measurement of the strength of relationship between two variables. The bigger the relationship, the more likely the finding has significance in practice.
Benchmark	A standard or point of reference against which things can be measured. Benchmarks may be externally or independently set or emerge 'normatively' – that is, what appears to be normal or average when lots of people engage over time.

Standard One – theory, design, plan

Is your ***theory of change*** attentive to the role of

- the organisational context (for example, resourcing, commitment to learning and quality);
- practitioners/managers (for example, lived/practice-based experience and levels of training);
- place and/or community-level influence (for example, pollution and climate change, levels of deprivation, local government); and
- levels of engagement at the ‘point of engagement’ (see [Outcomes Frameworks 2.1 and 3.0](#))?

Does your ***theory of change*** differentiate between short, intermediate and long-term outcomes for young people?

Does your ***theory of change*** include the following causes and effects within provision:

- dosage;
- activity type;
- activity fidelity/quality;
- young people’s skill level at ‘entry’; and
- skill/status change outcomes for individuals?

Have you compiled an ***evaluation plan*** with research questions and improvement goals that identify specific causes and effects named in the theory of change?

Have you compiled ***prior evidence*** (internal or external) that provides benchmarks i.e., answering questions about how ‘much’ of each cause is necessary to produce an effect? Can you anticipate different effects for young people who enter with lower or higher skills at baseline?

Do you have a ***clear design for your CQI cycle*** that will address improvement goals? Do you have a ***clear design for your evaluation*** that will address research questions about outcomes and/or impact?

Have you selected **measures** aligned to the theory of change, i.e., aligned to specific causes and effects named in the theory of change?

Do you have: a ***clear timeline*** for your CQI cycle and/or evaluation; ***clear selection/recruitment process*** for provision and participants; and ***clear procedures*** for informed consent, data collection, and data security (i.e., ethics)?

Considering all of the above, do you have the staff/volunteer capacity and other ***resources to implement*** your design for CQI and/or evaluation?

Standard Two – evidence about causes

Do handbooks, manuals, monitoring and quality procedures and staff training materials etc. ensure that provision can be ***delivered with fidelity*** in regard to dosage, provision type, quality, and young people's skill level on entry?

Are you keeping ***accurate records*** for the provision type, activity type, dosage/attendance, and demographics for each young person?

Have you selected a ***fidelity/quality practice measure*** and a method of collecting the data about practice (e.g., adult self- or external-observer; youth engagement survey)?

Have you selected an ***age-appropriate and aligned youth skill measure*** (i.e., aligned with the theory of change) to assess the selected or sampled group of young people's skills/needs at baseline?

Do you have a plan and resources to ***train and support staff and/or young people for data collection***?

Does your plan include converting the three types of data in this standard into ***CQI feedback in real time*** (within one or two weeks of collection)?

Standard Three – evidence about outcomes

Have you used your theory of change to anticipate the size of effect likely produced and then selected a youth outcome ***measure of sufficient sensitivity*** to capture the effects anticipated in the theory of change?

Have you selected a measure and procedures that are ***valid for calculating growth*** in simple models i.e., subtracting pre from post for each individual and then taking the average of the differences?

Do your ***results indicate a positive change*** in at least one of the provision's main outcomes from baseline to follow up?

Is there transparency in reporting ***the proportion of young people by type of change*** (increase, stable, decline in socio-emotional skills) and differences for important subgroups e.g., age groups, exposure to high or low fidelity/quality, ethnicity?

Have you provided details about the ***sample representativeness and data analytics*** procedures?

Standard Four – evidence about impact and replication

Have you gathered results that allow you to make an ***inference about growth, stability, or decline*** in at least one of the provision's main outcomes?

Do you have a ***comparison group*** and is there transparency about comparison group members' selection and participation in provision similar to the provision that is being evaluated?

Have you described ***sources of potential bias in the results*** such as unmeasured differences between the groups, or other weaknesses or limitations of the impact design employed?

Do your records, fidelity/quality, and baseline skill data indicate the ***proportion of young people targeted*** by provision (i.e., with lower skills on entry) ***that actually received*** the full provision as intended?

Have you ***integrated evidence about causes*** (dose, type, quality, baseline) to understand how different levels of provision lead to different outcomes for different young people?

Have you and/or external evaluators carried out two or more evaluations using a comparison group that ***replicate the same pattern*** of results?

Have you completed a ***cost analysis or break even analysis***?

THE CENTRE
FOR YOUTH
IMPACT

PART OF YMCA ENGLAND & WALES



Here for young people
Here for communities
Here for you

YMCA enables people to develop their full potential in mind, body and spirit. Inspired by, and faithful to, our Christian values, we create supportive, inclusive and energising communities, where young people can truly belong, contribute and thrive.

FAMILY & YOUTH WORK

HEALTH & WELLBEING

HOUSING

TRAINING & EDUCATION

SUPPORT & ADVICE