HOW TO RUN AN
„*Impact Evaluation*"

A Capacity Development Guideline

SPORT FOR DEVELOPMENT

Click the buttons to navigate through the document

Introduction

The CD Guideline 'Impact Evaluation' is part of the → *Sport for Development Toolkit Learning Lab*.

Do you want to run an impact evaluation in the field of 'Sport for Development'?

You want to ...

   ... measure possible impacts of your programme or intervention?
   ... apply scientific methods in your evaluation practice?

This document will support you. It explains the different steps to conduct an impact evaluation. Through inputs and self-learning exercises you will learn how to implement an impact evaluation with a sport-specific approach. This document consists of the following seven steps:

   1) Get a common understanding
   2) Define targets and indicators
   3) Design tools
   4) Select and train interviewers
   5) Select survey sample
   6) Collect data
   7) Analyse data & write reports

Target groups

Institutions and organisations interested in using sport-specific approaches for impact evaluation as well as S4D Instructors.[1]

General learning objective

You will be able to...

   • identify the different steps of an impact evaluation.
   • apply the provided material in your own working context.

Please note that all contents were developed in a working context of GIZ. Feel free to adapt and modify contents and scope according to your needs, e. g. regarding wordings, country examples.

[1]   For definitions of S4D Instructors and Coaches, see the → *GLOSSARY*.

SPORT FOR DEVELOPMENT

The structure of each step

Each step has
   • specific learning objectives and
   • the following components:

INPUTS
theoretical inputs including the most relevant information and related links

ASSESSMENTS
specific assessment questions, which you can use to test your knowledge gain regarding the inputs

SELF LEARNING
a preparation task for the upcoming step

Content

Step 1: Get a common understanding

INPUT: The Sustainable Development Goals
INPUT: Impact Evaluation
ASSESSMENT
SELF-LEARNING

Step 2: Define targets and indicators

INPUT: Frameworks
INPUT: Programme Objectives
INPUT: Indicator
INPUT: Using S4D competences as indicators
ASSESSMENT
SELF-LEARNING

Step 3: Design tools

INPUT: How can we measure indicators?
INPUT: Questionnaire
INPUT: Focus Groups Dicussions
INPUT: Mixed Methods Approach
ASSESSMENT
SELF-LEARNING

Step 4: Select and train interviewers

INPUT: Requirements and training content
ASSESSMENT
SELF-LEARNING

Step 5: Select survey sample

INPUT: Requirements for intervention and comparison/ control group
INPUT: Create information for the target group
ASSESSMENT
SELF-LEARNING

Step 6: Collect data

INPUT: Pilot testing and revision of quantitative instruments
INPUT: Data Collection Process
ASSESSMENT
SELF-LEARNING

Step 7: Analyse data & write reports

INPUT: Evaluation
INPUT: Data-Evaluation-Plan
ASSESSMENT
SELF-LEARNING

Learning objectives:

By the end of this step, you will be able to …

✓ … define an impact.
✓ … discuss the general concept of impact evaluation.
✓ … use a problem tree.

Step 1

Get a common understanding

INPUT
The Sports for Development Approach (S4D)

<<<< What is S4D >>>>

SPORT	SPORT PLUS[1]	PLUS SPORT	NO SPORT

Sport Development

Focusing SOLELY on sport objectives, such as the development of sport specific competences (techniques, tactics), competition etc.

Sport for Development

Using sport as core activity, adapted in various ways to achieve certain development objectives, e. g. healthy lifestyles.

Using sport as an entry point, means or additional component for development objectives, e. g. employability.

Development (without sport)

Other development activities and approaches that do NOT involve sport (e.g. specific education, peace building, social cohesion programmes) etc.

Sport for Development (S4D) refers to the intentional use of sport, physical activity and play to attain specific development objectives, including, most notably, the UN Sustainable Development Goals (SDGs).

The approach is based on the conviction that sport-based initiatives can be powerful and cost-effective means for achieving development objectives, such as education, health and gender and youth empowerment.

Sport for this purpose includes all forms of physical activity that contribute to physical and mental well-being and social interaction, such as play, recreation, organized or competitive sport, and indigenous sports and games.

The S4D approach aims at realizing the right of all members of society to participate in sport and it embodies the best values of sport while upholding the quality and integrity of the sport experience.[2]

The Sports for Development Approach (S4D)

1 The "Plus-Sport, Sport-Model" was adapted from Fred Coalter (see Coalter & Taylor, 2010, p. 1)
2 Our definition contains parts of the following documents: Sport, Recreation and Play (UNICEF, 2004, S. 1); Harnessing the Power of Sport for Development and Peace (SDPIEG, 2008, S. 1); Transforming our World, Agenda 2030 (UN A/RES/70/1, Art. 37); Sport for Development and Peace, Towards achieving the MDGs (UN, 2003, p. 5); What is Sport for Development and Peace (RTP, p. 1); Introduction to Sport for Development and Peace (p. 3);

INPUT
The Sustainable Development Goals

The 17 Sustainable Development Goals (SDGs) have been implemented by all UN-Members in 2015 through the 'Agenda 2030 for Sustainable Development'. The SDGs aim to reduce poverty, protect our planet and achieve peaceful living. Sport has been recognized as a contributor to achieve those goals. The Commonwealth Secretariat has created several indicators to help evaluate physical activity programmes and measure their contribution to the SDGs. The indicators and more information can be found on the → *Commonwealth website*.

The Sustainable Development Goals (United Nations, 2019)

INPUT
Impact Evaluation

Internationally, the approach that sport can contribute towards the Sustainable Development Goals (SDGs) has received increasing attention over the last decade.

The overall objective of impact evaluation in S4D is to measure the impact of sport on selected SDGs. Therefore, the SGD targets need to be concretised into measurable indicators.

For example: Through a long-term and regular participation in the S4D courses children and youth acquire competences which are beneficial for their employability.

What are impacts?

Definition: *Impacts are intended and unintended, positive or negative changes of a state, structures, processes or behaviour due to an intervention. (GIZ 2018)*

- If a causal relationship between the performance and the effect of an intervention can clearly and plausibly be demonstrated, the observed change can be declared as an impact of the programme.
- Proving that the S4D approach is actually working (in regards to meeting its own objectives) is a major part of our legitimisation as a programme!

What is an impact evaluation?

An impact evaluation is an analysis that examines the medium and long-term outcomes and impacts of an intervention, and looks at the extent to which the observed results (changes) were caused by and can be traced back to the intervention.

A causal relationship between an intervention and the results can either be clearly demonstrated using quantitative methods (attribution) or plausibly shown using qualitative methods (association). Independent of whether an evaluation is based on quantitative or qualitative methods, it always examines what would have happened without the intervention. Analysis of this 'counterfactual situation' enables different or alternative development paths to be assessed (i.e. a comparison of the initial situation with the post-intervention situation) (GIZ 2018).

What are general challenges with measuring impacts?

- It is likely that many of the involved persons can somehow see and feel a change through applying S4D with children & youth. However, this change cannot be quantified yet.
- There are other influencing factors that can cause a relationship (alternative hypotheses):
  - Environmental impact, other sport programmes and interventions
  - Time of measurement
  - Selection bias, panel mortality
- Therefore, colliding factors need to be considered and controlled.

Problem: S4D Competences like *respect, empathy, self-confidence* are not directly observable and therefore need to be measured *indirectly*.

The Problem Tree

The problem tree is a tool to analyse known effects and causes of a specific problem. It helps to understand the complexity of a problem and the context in which an intervention is taking place. The problem tree is not only a tool used in the evaluation practice but is also applied as a project planning tool.

Example:

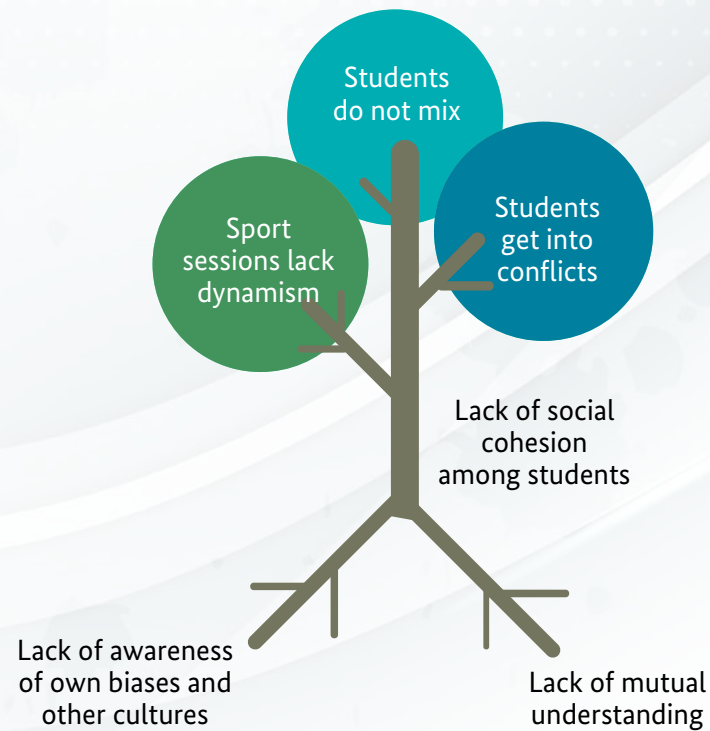**Branches:** What **effects** does the problem have?

Trunk: What is the **focal problem?**

Roots: What are the **causes?**

Students do not mix

Sport sessions lack dynamism

Students get into conflicts

Lack of social cohesion among students

Lack of awareness of own biases and other cultures

Lack of mutual understanding

Resources to use:

→ *Guideline "The SDGs and Sport"*
→ *The Commonwealth - Measuring the contribution of sport to the Sustainable Development Goals*
→ *United Nations – The SDGs*

SPORT FOR DEVELOPMENT

ASSESSMENT

1. What is an impact?

A: An intended change only
B: A change of state, behaviour, or structure due to an intervention
C: Negative and positive changes

2. What is an impact evaluation?

A: An analysis of outputs
B: Impacts, focal problem, causes
C: An analysis of medium and long-term effects

3. Which three elements does a problem tree contain?

A: Effects, focal problem, context
B: Impacts, focal problem, causes
C: Effects, focal problem, causes

SELF-LEARNING

Task: How does your problem tree look like?

1. What are the **focal problems** in your community, school or environment?

2. What are the **causes** of those focal problems?

3. What are the **effects** of the focal problems?

SPORT FOR
DEVELOPMENT
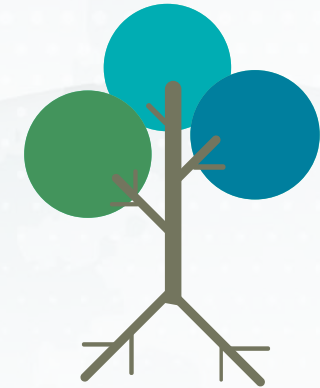
Learning objectives

By the end of this step, you will be able to …

✓ … predict which frameworks can be used for a theory-based approach.
✓ … formulate programme objectives.
✓ … explain the relevance of indicators and how the indicator frameworks
on the Resource- Toolkit-Website can be used in this context.
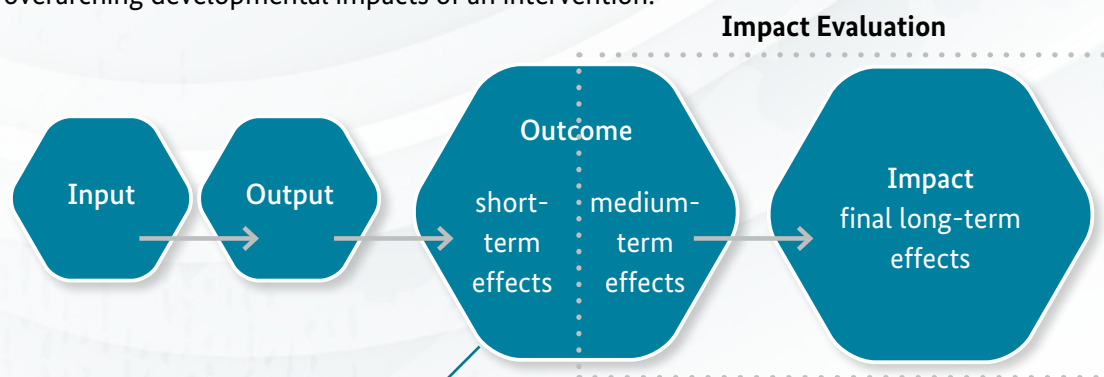
Step 2
Define targets and indicators

INPUT
Frameworks

Preconditions for impact evaluation:

To enable *robust* proof of impact, a theory-based evaluation approach is needed, which uses one or several of the following elements: **Result Chain/ Theory of Change/ Result Model.** These are used to visualize cause-and-effect relationships (from inputs to outputs to outcomes to impacts) and possibly explains mechanisms of action. Which one you use depends on your preferences and the fit to your evaluation object.

1. Result Chain

Depicts the most important levels of impact evaluations: the medium-term and long-term, overarching developmental impacts of an intervention:

Impact Evaluation

Input → Output → Outcome (short-term effects : medium-term effects) → Impact final long-term effects

The Result Chain (Caspari & Barbu 2008)

Definitions

- **Input:** Resources that are needed for the activity, e.g., financial, human, or material resources.
- **Activities:** Actions and interventions to achieve specific outputs (often not depicted).
- **Outputs:** Products produced, or services provided by a programme.
  - > Examples: Classes held, football practices offered, learning materials or films provided.
  - > In short, outputs can include services, activities, or products.
- **Outcomes:** Results within the target group (e.g., knowledge & skills gained); outcomes can be further subdivided:
  - > The short-term outcome level is reached when we observe *changes in skill levels and/or attitudes within the target group.*
  - > The medium-term outcome level is reached when we observe changes in decision-making and behaviour and the members of the target group are able to improve their living conditions *(target group's circumstances change).*
- **Impact:** Long-term results at the societal level (e.g., decline in unemployment), developmental effects (SDGs).

2. Theory of Change

"A theory of change is the starting point of strategy development and the basic structure for an M&E system" (Elsemann, Hebel, Jäger & Daraspe, p.17).

"A theory of change defines the organization's goals and maps the necessary changes that must take place at both target group and stakeholder levels, as well as the external factors. It is the basis for understanding what information is required and what data should be collected to measure the organization's progress towards its goals" (Elsemann et al., p.11).

Inputs (resources)	Output (activities)	Outcomes (results on target group level)	Impacts (results on societal level)
What we invest	**What we offer, people we reach**	**What we achieve within the target group**	**What we want to contribute to**
• Number of staff • Amount of money • Material • Etc.	Activities were carried out • Number of S4D workshops • Number of S4D training sessions	Changes in knowledge & competence gain • Coaches know more about S4D • Youths know more about training topics (e.g., self-confidence, communication) • Gained social/ personal competences	Changes in society • Change in youth unemployment rate • Change in access for persons with disabilities in the region
	Target groups are reached (disaggregated by gender) • Number of coaches trained • Number of youths participating • Number of persons with disabilities participating	Changes in attitude & behaviour • Coaches are able to implement S4D training sessions including persons with disabilities • Youth act in solidarity with each other • Youth feel more confident	
	Target groups accept offers (disaggregated by gender) • Number of coaches satisfied, taking part • Number of children with/ without disabilities taking part	Changes in socio-economic status/ living conditions • Participants are qualified coaches • Youth found a job • Persons with disabilities have access to sport sessions	

3. Results-based Model

GIZ's results-based models are systemic, not linear relations. They incorporate the elements of impacts, outcomes and outputs used in the OECD-DAC's definition of results.

The results model depicts the causal links and forms the basis for formulating the hypotheses – i.e. the assumed causal relationship between an intervention and a result. Results are understood as the changes that can be attributed to a project or object of an evaluation. The mere occurrence of a change is not enough to designate it as a result of the project, even if the change was planned and intended. The observed change can only be deemed to be a project result if a **causal link** can be clearly or plausibly substantiated. Results may be intended or unintended, expected or unexpected, positive or negative. Objectives, on the other hand, are merely intended positive results (GIZ 2018).

INPUT
Programme Objectives

What are programme objectives?

- A programme's intended results, which contribute to improving physical, financial, institutional, social, environmental or other conditions for people, groups, organizations or elements of the broader society.
- Programme objectives at the societal level (impacts) describe the long-term results that are achieved or influenced by the programme.
  > These societal-level objectives are closely linked to the organization's vision.

Why programme objectives?

✓ Provide orientation
✓ Establish a basis for impact evaluation
✓ Motivate the programme staff
✓ Convince third parties of the programme's values

Be careful with the term 'impact'!

Developments at the societal level are determined by many factors. Demonstrating a causal relationship between a programme and impacts can often be difficult, and may even be impossible, especially since impacts typically appear only after considerable time.

> Programme-specific objectives at the societal level should thus be formulated cautiously ("The programme will contribute to...").

Distinction between direct and indirect contribution

"A *direct contribution* may be defined and identified where a sport-based approach has an identifiable impact, independent of other factors, by, for example, [...] making a direct contribution to the achievement of equal opportunities for women through specific employment practices.

An *indirect contribution* may be achieved alongside other activities beyond the realm of sport. [...] Examples may include the use of evidence on the contribution of sport-based activities to specific aspects of public health education" (Lindsey & Chapman, 2017, S. 35).

INPUT
Indicator

Definition
An indicator is a *pre-defined variable* which helps to identify (in)direct differences in quality and/or quantity within a defined period of time.

- *Unit of measure* helps to judge if an intervention was successful or not.
- Complex problems are *simplified and reduced* to an observable dimension.

An indicator can be thought of as a sign that something has changed. Thus, yellow leaves on the trees are an indicator that autumn is coming, and rustling leaves are an indicator of the presence of wind.

- Indicators are quantifiable variables and *indicate*. Indicators should be **SMART** (Doran, 1981):
- **S**pecific: Programme objectives need to be formulated clearly. Try to be as precise and transparent as possible when defining objectives, thereby enabling third parties to under- stand them, too.
- **M**easurable: Objectives must be subject to quantifiable description.
- **A**ccepted: Programme objectives need to be accepted by stakeholders.
- **R**ealistic: It must be possible to actually attain the formulated programme objectives ('doable').
- **T**ime-bound: When formulating programme objectives, it can be challenging to deter- mine a fixed point in time for achieving them. Nonetheless, you should be able to provide

an estimated schedule for achieving programme objectives. Objectives can be reached during the programme's operation or later, but it is important to have a schedule to keep the programme on track.

Indicator ≠ Indicator

- **Results-level indicators** (outcome and impact indicators) help you determine whether and to what degree your programme is achieving results.
- **Output-level indicators**: Although outputs are not themselves results, they are the *prerequisites for achieving results* on higher levels (only information at the beginning).
  > *Be careful: The fact that output-level indicators are easy to develop and monitor some- times leads managers to focus primarily on these even later in the programme period.*
- **Input-level indicators** are also relevant, since they …
  … provide information on the resources going into the programme, and
  … enable conclusions to be drawn regarding the programme's efficiency and efficacy.
- **Quantitative indicators** are numerical. (Elsemann, Hebel, Jäger & Daraspe, o.J.)
  They are used to examine the scale or prevalence of a change, e.g. the number of people who attended a training course.
- **Qualitative indicators** are descriptive (Elsemann et al., o.J.)
  They tell you about the character or quality of something, using words, video, photos or diagrams.

INPUT
Using S4D competences as indicators

What can the trainer observe?

• It is likely that trainers can somehow see or feel a change through applying S4D with children and youth, but how can we measure this effect?
> We need to translate the SDGs into quantifiable indicators!

We cannot measure life skills or competences directly.
> Therefore, we designed a *framework to systematically cluster competences.*
> This helps to evaluate our impact on children and youth.
> In these frameworks we build *the bridge* between our S4D activities, competences, SDGs & impact.

S4D Competence Framework
The competence frameworks cluster S4D competences that children and youth can gain in the field of S4D. In the competence frameworks, SDG-related competences are streamlined, including specific competences in the areas of *self-competence, social competence, methodical/ strategic competence,* and *sport-specific competence.* In line with other theoretical systems and to explain change on an individual level, the competences are divided into three levels: Recognising, Assessing and Acting.

Competences Children and Youth gain in Sport for GENDER EQUALITY Programmes (13-16 years)	
Self-competence Children and youth are able to …	**Self-confidence and Trust** … recognise the importance to turn to teammates or trainers when facing discrimination or violence based on gender ond/or sexual orientation. … reflect on situations where they were confident regarding their gender identities and/or sexual orientation despite. … ensure privacy when changing into sports gear.
	Motivation … recognise the importance of role models (such as female/male as well as LGBT+ athletes/coaches/teachers). … demonstrate their motivation to realize their own objectives, independent of stereotypes based on gender and/or sexual orientation. … make concrete plans to become sport role models.

Example of a S4D competence framework related to SDG 5 Gender Equality

S4D Impact Indicator Framework

The S4D Impact Indicator Frameworks advocate for a systematic and specific approach to measure the contribution of sport, physical education, and physical activity to the SDGs. The impact indicator frameworks are based on the S4D competence frameworks.

✓ We translated *competences* into *indicators*
✓ The competences children and youth acquire in our S4D activities are the indicators for the impact evaluation
✓ It will finally be possible to proof the links between SDGs and the S4D activities
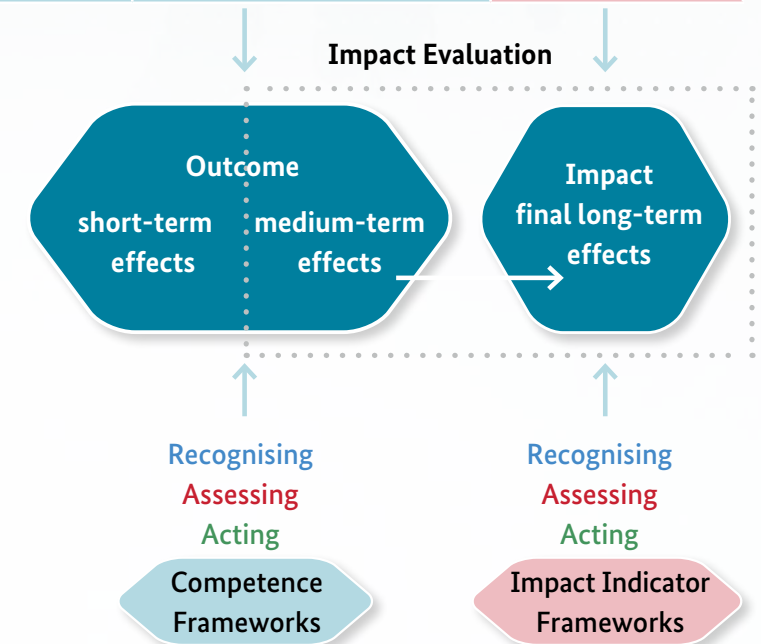✓ The framework helps a lot to look at the whole picture

(Possible) hypothesis:

If children/youth regularly take part in S4D training sessions involving S4D activities focusing on an SDG they may gain S4D specific competences contributing to SDG targets.

The following example of an Impact Indicator Framework is related to the SDG 5 Gender Equality and shows how specific competences from the framework can be translated into indicators to measure to what extent the S4D Approach contributes to SDG Targets.

General Dimension: S4D Competence Cluster	Sub-Dimension: S4D Competences	S4D Sub-Indicator: Selected S4D specific competences	S4D Main Indicators:	SDG Target:
		XY % of children/youth...	XY% of children/youth...	Target 5.2: Eliminate all forms of violence against women and girls in the public and private spheres, including trafficking and other types of exploitation
Self-competence	Self Confidence and Trust	... recognise the importance to turn to team mates or trainers when facing discrimination or violence based on gender and/or sexual orientation	... explain the basic rights of women, girls, LGBT+, including their right to freedom from exploitation and violence, as well as their reproductive rights. ... reflect about all forms of violence against girls and women (incl. trafficking and other forms of exploitation)	
	Resilience	... acquire information on strategies to recover from difficulties in life created by discriminations based on sexual orientation and/or gender. ... stick to their goals despite ongoing inequalities based on sexual orientation and/or gender.		
	Creativity	... acquire information about safe spaces to go in for sports.		

Placing the competence framework and the impact indicator frameworks in the Result Chain (own illustration)

Impact Evaluation

Outcome
short-term effects — medium-term effects

Impact
final long-term effects

Recognising
Assessing
Acting

Competence Frameworks

Recognising
Assessing
Acting

Impact Indicator Frameworks

✓ The specific competence is the indicator for the difference between S4D participation and no S4D participation.
✓ A survey is used as a test if this hypothesis can be maintained or is false.
✓ For generating a survey, all our defined *indicators* need to be transferred into *questions* that we can ask children and youth.

Competences → Indicators → Questions
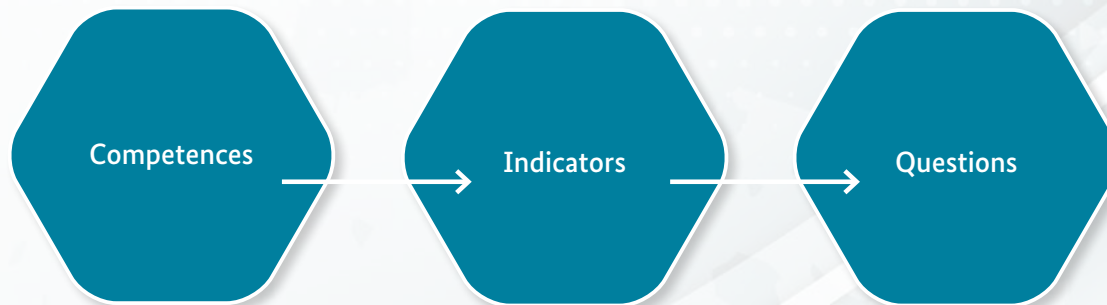
Literature

- Caspari, A., & Barbu, R. (2008). *Wirkungsevaluierungen. Zum Stand der internationalen Diskussion und dessen Relevanz für die Evaluierung der deutschen Entwicklungszusammenarbeit* (CEval Arbeitspapier 16). Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung.
- Doran, G. T. (1981). There's a S.M.A.R.T. way to write management's goals and objectives. *Management Review.* 70 (11): 35–36.
- Elsemann, K., Hebel, M., Jäger, L., & Daraspe, C. (o.J.). *Monitoring and Evaluation in Sport for Development: streetfootballworld.* Accessed March 13, 2020 → *here*
- Lindsey, I. & Chapman, T. (2017). *Enhancing the contribution of sport to the sustainable development goals.* London: Commonwealth Secretariat.

Resources to use

→ *Manual Streetfootballworld Monitoring and Evaluation*
- Commonwealth Secretariat (2019): →*Sport and SDGs Indicator Framework*
→ *Competence Frameworks* (under: S4D Competences/ Life Skills – Children and Youth Competences)
→ *Impact Indicator Frameworks* (under: M&E in the field of S4D, Impact Indicator Frameworks are provided for SDGs 3,4,5,8,10,13 and 16 in four different languages)
- Building Bridges between activities and impact-→ *Document*

ASSESSMENT

Task 1: Define your impact-oriented objectives

✓ What do you want to **achieve**, especially for the target group?
✓ How do you decide that your work has been **successful**?
✓ How do you determine what concrete **results** have been achieved within the target group?
✓ Which **target group** do you want to reach?
✓ What **changes** should the programme bring about within the target group?
✓ What objectives should the programme contribute to pursuing at **societal level**?

CHECKLIST programme objectives!

Outcome objectives	Impact objectives
• **Benefits** for the target group produced by the programme • **Changes** within the target group as a result of the programme's work • Success can be attributed	• **Long-term effects** on societal level the programme helps to achieve/influence. • The programme is contributing to [name your objective] • Results cannot be attributed exclusively to the programme's work
✓ Specify your target group ✓ Use active verbs ✓ Ask yourself: What new opportunities have emerged through participation? How has the living situation changed? ✓ Express the objective positively	✓ Formulated more abstractly ✓ Relate to the society as a whole ✓ Time frame not needed ✓ Convert societal problems into positive statements

Task 2: Formulate your own indicators

STEP 1: Collect ideas! Starting point are your objectives!

Objective	Dimension	Indicator
After programme participation, youth have secured apprenticeships (directly verifiable)	countable	Number of youth that secured a job within six months after participation in the programme
Youth gained better job-application skills (not directly verifiable)	countable	Number of participants in training sessions
		Number of youth hired after applying for a job
	countable	Youth are able to write a good job application
		Youth have a clear career perspective

STEP 2: Structure and refine ideas

> Some programme objectives can be captured using a single indicator, for instance using quantitative characteristics.
> More complex programme objectives will usually require more indicators, often drawing on qualitative and quantitative indicators simultaneously.

STEP 3: Check your indicators

Are they **SMART**? specific, measurable, accepted, realistic and time-limited
Your indicators should cover the following questions:
> What information do you need to determine whether the programme participants have undergone the desired transformation?
> What information will help you determine how to improve or adapt your programme?
> What would you need in order to tell that something is going wrong?

> Try to find a **good mix of indicators** that illuminates both the quantitative and qualitative aspects of the programme.

STEP 4 Selection of indicators

• It's time to set priorities!
• The goal is to have a small but meaningful set of indicators!

> CHECKLIST for indicators

SELF-LEARNING

Task: Develop your own Theory of Change

Inputs (resources)	Output (activities)	Outcomes (results on target group level)	Impacts (results on societal level)
What we invest	**What we offer, people we reach**	**What we achieve within the target group**	**What we want to contribute to**
…	Activities were carried out …	Changes in knowledge & competence gain …	Changes in society …
	Target groups are reached …	Changes in attitude & behaviour …	
	Target groups accept offers …	Changes in socio-economic status/ living conditions …	

Step 3
Design tools

Learning objectives

By the end of this step, you will be able to …

✓ … set up a survey.
✓ … differentiate between the different evaluation tools as well as quantitative and qualitative surveys.
✓ … design a questionnaire and a focus group discussion and discuss validated instruments that have already been used in the S4D context.
✓ … connect instruments with the mixed method approach.

INPUT
How can we measure indicators?

Surveys …

… can measure characteristics, opinions, attitudes, or beliefs.
… cannot measure change in behaviour directly, but use the concept of self-assessment.
… should consider the social context and background (age, language, culture, education, environment) of the target group.

	Quantitative	Qualitative
Procedure	Deductive, 'to measure': counting, collecting numbers (e.g., demographic data)	Inductive, 'to understand the meaning': open questions, open tasks
Product	Hard, replicable data	Soft, realistic data
Process	Static	Dynamic
Perspective	External perspective of the researcher	Internal perspective of the people concerned
Interest	Explanation of causal relationships, generalisations	Exploration of the living environment, interactions, and relations
Evaluation tools	✓ Standardized questionnaire	✓ Personal interview: Semi-structured or non-directive interviewing ✓ Participant observation ✓ Focus Group Discussion ✓ Story telling ✓ Diaries ✓ Blogs ✓ Photo or video monitoring ✓ Performing arts ✓ Painting and drawings

INPUT
Questionnaire

> **Check the length of your questionnaire. Some participants will maybe quit if it takes too long to answer.**

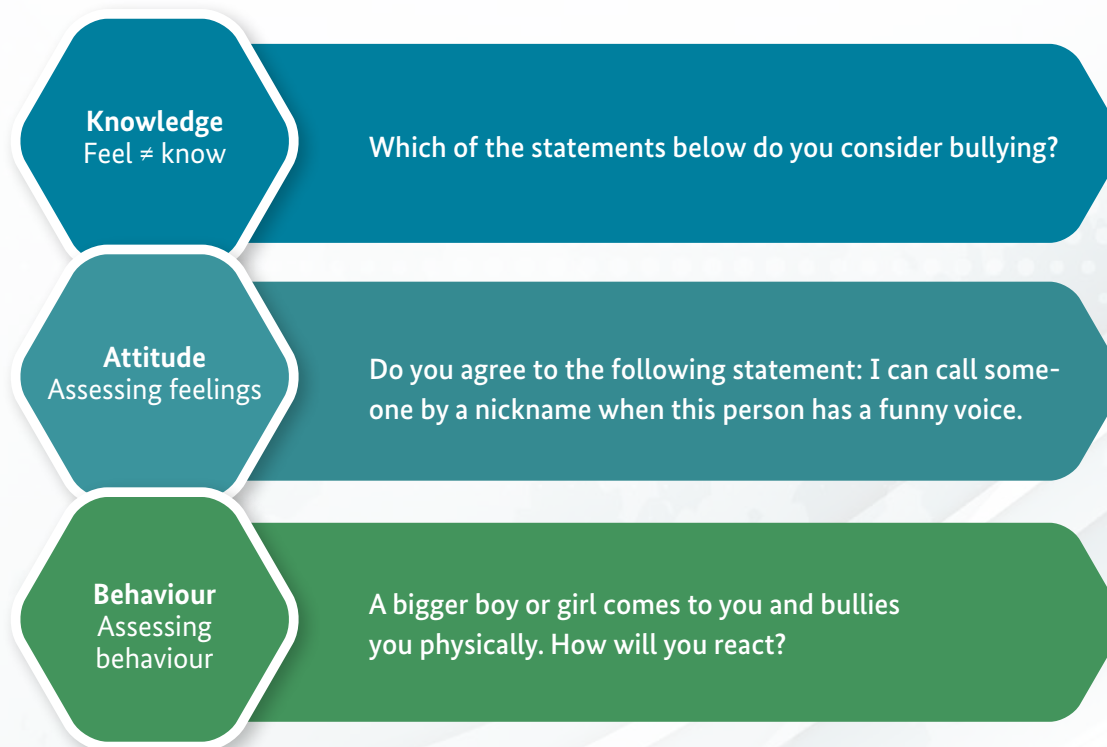**Steps to take to develop a questionnaire**

- A *questionnaire* is probably the most used *scientific method* to *systematically gather desired information*.
  - High standardisation of questions and survey situation leads to comparability of results.
  - 'Closed' answers allow easy and fast analysis.
  - Questions about knowledge, attitudes, behaviours or characteristics are possible.
  - Idea: The respondent is the 'expert'.
- The questionnaire is our link to measure the *indicators*.
- You can either develop a questionnaire yourself, specific to your context and target group, or use a validated instrument (see below).
- Questionnaires can be paper based or digital, with the latter showing many advantages during data collection and analysis.

Finding the right questions

- *Always have the target group (respondents) in mind!*
- Use easy words and short sentences.
- Be as precise as possible.
- Try not to provoke a specific answer.
- Formulate neutral.
- Only ask one question each at a time.
- Do not use double negations.

How could a systematic approach measuring the impact in S4D look like?

- ***The Basis: Competence Framework*** (have a look at the resources)
- Can a youth (~14yrs) answer this question?
- Do the questions depict the indicator?
- Is the category of knowledge, attitude, behaviour fitting?
- What do we like to find out by asking this question?

SPORT FOR DEVELOPMENT

Knowledge
Feel ≠ know

Which of the statements below do you consider bullying?

Attitude
Assessing feelings

Do you agree to the following statement: I can call someone by a nickname when this person has a funny voice.

Behaviour
Assessing behaviour

A bigger boy or girl comes to you and bullies you physically. How will you react?

Examples from Iraq on how to assess change in knowledge, attitude and behaviour

Questionnaire examples from Iraq and Western Balkans can be found on the
→ *Toolkit-Website*.

Validated Instruments already used in S4D

Validated instruments are instruments that have undergone several statistical tests and that were reviewed by other experts to determine whether a certain questionnaire assesses a certain construct. If you are interested in more details, look up the different categories of validity: content-related evidence, criterion-related evidence and construct-related evidence.

Example 1: Self Efficacy Scale

- Self-efficacy beliefs determine how people *feel, think, behave and motivate* themselves (Bandura 1994).
- Self-efficacy is "an individual's *belief in her/his ability* to plan and perform a task, to achieve a particular outcome [and] to address difficult issues" (Coalter 2013, p.60).
- In psychology and education, self-efficacy has proven to be a consistent predictor of behavioural outcomes.
- The most effective way of creating strong self-efficacy beliefs is through *experience*.
    > *S4D programs might be an effective medium since the emphasis lies on practice, competences development, on dealing with and learning from defeat and on social climate.*
→ *Self Efficacy Scale Questionnaire*

Don't change validated instruments!
Use them as they are **or design a completely new instrument!**
The development of a new instrument is sometimes better, because you can tailor it to the programme (content, objectives, SDG targets etc.)

INPUT
Focus Group Discussion

Example 2: PYD (Positive Youth Development) Questionnaire

- "PYD is an *umbrella term* rather than a singular construct or theory. It is used in both research and practice and tends to focus on the adolescent period of development. Adolescence is marked by significant changes in emotional, cognitive, social, and physical domains." (Holt et al. 2019, p. 342)
- Through sport youth can acquire competences, assets, values, and life skills that will have a positive influence on their overall development.
- The Five Cs Model of PYD (Lerner et al. 2003) belongs to the most empirically supported frameworks and emphasizes the strengths of youth as well as enables them to be seen as resources to be developed.

→ *Positive Youth Development Questionnaire*

A Focus Group Discussion (FGD) is a qualitative research method that groups together four to twelve people to discuss a certain topic. FGD focus on the interaction of the participants and aims to acquire information about thought processes, feelings, and experiences. The moderator functions as a guide and leads the group interaction. A focus group discussion can also integrate activities and games to create group dynamics. Researchers using focus group discussions are interested in how individuals discuss a particular topic in a group. The focus is more on the group dynamics and less on the individual. Reactions of participants to opinions of other participants and the development of new opinions are of interest. Depending on the age of the participants, the moderator asks a broad range of questions to generate a discussion. In the case of children and youth, more guidance and structure are needed by the moderator.

FGD is often combined with standardized quantitative surveys to understand and examine certain topics/correlations more in-depth.

- ✓ Group discussion (4 -12 people) to discuss a certain topic.
- ✓ Acquire information about thought processes, feelings & experiences.
- ✓ Group interaction.
- ✓ Moderator as a guide.
- ✓ Integrate activities & games.
- ✓ The composition of the group affects the dynamics of the discussion.

Steps to take to develop guidelines for an FGD

- Formulate an introduction for the participants, including information about the schedule and the purpose.
- Ensure anonymity of participants and the protection of personal data.
- Get permission to record the Focus Group Discussion.
- Select a moderator who is independent and is not associated by the participants with your project (e.g. not a coach) but has expert knowledge on your research topic.
- Plan a warm-up game or activity.
- Create open questions for your FDG:
    - *Always have the target group (respondents) in mind!*
    - Use easy words and short sentences.
    - Be as precise as possible.
    - Try not to provoke a specific answer.
    - Formulate neutral questions.
    - Only ask one question each.
    - Do not use double negations.
- Include activities/interactions, if desired.
- Focus group discussions should be recorded to transcribe and analyse them later.
- Plan your material: observation sheet, voice recorder, pens, camera, etc..
- Plan the seating arrangement.
> Suggested seating arrangement:

A Focus Group Discussion is not an Interview! The objective is to initiate a discussion.
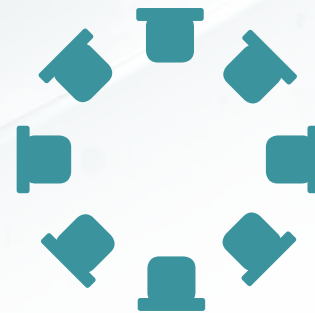
Implementation of an FGD

- Set up an initial stimulus that initiates a discussion.
- The style of your moderation varies depending on how the discussion is structured, the goal of the FGD, the topic(s) to be discussed, and the participants. You can either play a more active role in steering the discussion or withdraw into an observer role.
- In social sciences, an FGD ends when its immanent potential has been exhausted and no new insights are generated by asking. When conducting an FGD with children and youth, be aware that compared to adults their concentration decreases much faster. An FGD with children and youth should not be longer than 1 – 1.5 hours.

INPUT
Mixed Methods Approach

The mixed method approach involves collecting and analysing both quantitative and qualitative data:

Quantitative data	Qualitative data
✓ Closed-end information that undergoes statistical analyses ✓ Numerical representation ✓ Answers the question of whether effects have occurred and to what extent	✓ More subjective and open-ended ✓ Describe examples of impact at the target group level ('zoom in') ✓ Answers questions such as why and how effects have occurred ✓ Focus on learning process (rather than accountability)

How can the methodologies be mixed to provide a more thorough understanding of a research problem?

- Data can be collected using a quantitative data instrument, followed by interviewing participants to learn more about some of the survey responses, which provides a more thorough understanding of the results.
- Interviews can be conducted to explore how individuals describe or feel about a particular topic. Afterwards, the information can be used to develop a quantitative survey.

- The researcher may be planning to use quantitative methods to assess the impact of a particular intervention. He or she may conduct interviews to better recruit appropriate participants.
- The researcher uses focus group discussions to collect information regarding a certain topic and then uses a quantitative survey with a larger group to validate the responses of the FGD.

Requirements/Preconditions

Quantitative Methods (Standardized Questionnaire)	Qualitative Methods (Focus Group Discussion)
✓ The questionnaire needs to be tested before (pilot testing) ✓ Participants should be able to read ✓ The participants should not be too young (12+); otherwise, they will need support filling out the questionnaire (objective person) ✓ At least one person should be available for further questions/support during the data collection ✓ The same participants (intervention AND comparison/control group) must fill out the same questionnaire (at least two times)	✓ Size between 4 to 12 people, with one moderator as interviewer ✓ Discusses a certain topic ✓ Acquires information about the reflection, feelings & experiences from the participants about a particular topic ✓ Explicitly uses group interaction to produce data & insights ✓ Requires the moderator to guide participants to openly discuss the selected topics ✓ Has participants responding to both the interviewer and to each other ✓ Should be tape-recorded & transcribed

Literature

• Bandura, A. (1994). Self-efficacy. In R. J. Corsini (Ed.), *Encyclopedia of psychology* (2nd ed., Vol. 3, pp. 368-369). New York, NY: Wiley.
• Coalter, F. (2013). *Sport for Development. What game are we playing?* Routledge.
• Holt, N. L., Deal, C. J., Pankow, K., Pynn, S.R., & Jørgensen H. (2019). SDP and positive youth development. In H. Collison, S.C. Darnwell, R. Giulianotti & D. P. Howe (Ed.), *Routledge Handbook of Sport for Development and Peace* (p.341-351). Routledge.
• Lerner, R. M., Dowling, E. M., & Anderson, P. M. (2003). *Positive youth development.* Thriving as the basis of personhood and civil society. *Applied Development Science,* 7(3), 172-180
• Stockmann, R. (Hrsg.). (2007). Handbuch zur Evaluation. *Eine praktische Handlungsanleitung.* Waxmann.

Resources to use

• Guideline Focus-Group Discussion: → *FGD Guideline Namibia 2015*
• Guideline Focus-Group Discussion: → *Step-By-Step Guide for Focus Group Discussions*
→ *Competence Frameworks* (under: S4D Competences/ Life Skills – Children and Youth Competences)
→ *Self Efficacy Scale Questionnaire*
→ *Positive Youth Development Questionnaire*
→ *Positive Youth Development Tool Book*

ASSESSMENT

Name four different evaluation tools:

1. _____
2. _____
3. _____
4. _____

Where is the difference between a qualitative and a quantitative approach? Fill out the table:

	Quantitative	Qualitative
Procedure		
Product		
Process		
Perspective		
Interest		

Name two validated instruments to conduct a standardized questionnaire:

1. _____
2. _____

SELF-LEARNING

Task: Choose your methodical approach and design your survey tool.

Learning objectives

By the end of this step, you will be able to …

✓ … select qualified interviewers.
✓ … structure an interviewer training.

Step 4
Select and train interviewers

INPUT
Requirements and training content

When conducting a standardized survey among children and youth, the participants can either fill out the survey directly themselves (self-assessment) or interviewers can interview the children with the predefined, standardized survey, and fill out the answers. Especially with younger children, there is a certain advantage to working with interviewers, due to possible (digital) illiteracy of younger children. Working with interviewers is also helpful to prevent questions from not being understood.

When relying on interviewers, you need to make sure that they are adequately trained and qualified to handle situations such as unexpected participant behaviour. They need to know how to construct and deliver questions as well as how to handle sensitive topics. The interviewer and you as a researcher need to be aware of the consequences of the interviewer's behaviour during the interview situation and the possibility to influence a certain answer behaviour in the participant. Therefore, interviewers need to be well selected and trained.

Attributes of qualified interviewers

✓ **Appropriate age:** younger interviewers may be trusted more by children and youth participants.
✓ **Even gender proportion:** however, watch out for how the interviewers' gender may influence the interview situation. In some societies it may not be appropriate for a male interviewer to interview girls or women. Sometimes female participants might not feel they are in a safe and trustworthy environment with a male interviewer.
✓ **Independence:** interviewers should not know the participants before and vice versa (no S4D coaches!).
✓ **Knowledge of your intervention:** the interviewer should understand your activities, project concept and aims.
✓ **Commitment:** switching interviewers during an ongoing data collection is complicated. Therefore, they should commit for the whole data collection phase. You can ensure this by adequate renumeration.
✓ **Experience:** previous experience in data collection and standardized interviewing is an asset but not a requirement.

Interviewer training (1 day):

✓ Explain the context and aim of your evaluation.
✓ Describe the data collection process and logistics.
✓ Go through the entire questionnaire and discuss the questions. It is essential that the interviewers themselves understand the questions, as they need to explain them to the participants. However, interviewers need to stick to the questionnaire!
✓ If you use digital appliances such as smartphones or tablets, train the interviewers in their use and maintenance.
✓ Don't forget to make interviewers aware of data protection guidelines.
✓ Establish a code of conduct.
✓ Discuss sensitive questions and appropriate behaviour in such cases.
✓ If you are interviewing children or youth, invite an expert (coach, teacher, psychologist, researcher) who is experienced in working/doing research with children or youth. Social desirability (the tendency to answer in a manner that is viewed favourable by others) is a major factor, especially among children and youth during interviews.
✓ Teach the interviewers about the effects of nonverbal communication, such as body language and voice.
✓ Recreate interview situations and let the interviewers practice (role play). If possible, have them practice with volunteers similar to your targeted participants (but not the participants themselves!). Observe these practice interviews closely and discuss them together.
✓ Conduct frequent quality checks during the data collection.
✓ If you conduct a survey twice (e.g., before and after an intervention) consider a refreshment training.

> You can also provide the interviewers with a written guidance booklet.

Resources to use

• GESIS (2016) → *StieglerBiedinger_Interviewer_Skills_and_Training.pdf (gesis.org)*

ASSESSMENT

True or false?

1. It is beneficial to the study to use interviewers who are already familiar with the participants in order to build trust.

    True ◯   **False** ◯

2. Gender and age of the interviewers can influence the interview situation.

    True ◯   **False** ◯

3. Social desirability is the tendency to answer questions the opposite way a respondent feels.

    True ◯   **False** ◯

4. An interviewer training should not take longer than 3 hours.

    True ◯   **False** ◯

5. When using interviewers, a code of conduct is needed.

    True ◯   **False** ◯

6. Data protection is something that needs to be discussed with interviewers.

    True ◯   **False** ◯

SELF-LEARNING

Develop the appropriate criteria for interviewer selection that fits your study.

Think about the following questions:

- *Who are the participants in my survey?*
- *How can I make the interview situation comfortable for them?*
- *Is my evaluation topic sensitive? Do I have triggering questions in my survey?*
- *What expert knowledge is required to conduct the survey?*
- *What attributes does the perfect interviewer for my study have?*

Plan your interviewer training:

Topic	Who?	Timeframe

Learning objectives

By the end of this step, you will be able to …

✓ … draw a survey sample out of the target group.
✓ … differentiate between an intervention, comparison, and control group.
✓ … assess preconditions and challenges of a randomized control trial.

Step 5

Select survey sample

INPUT
Requirements for intervention and comparison/ control group

Intervention and comparison/ control group

To evaluate to what extent observed results (changes) were caused by and can be traced back to your project intervention, you need to establish an intervention and a control/comparison group. The intervention group consists of persons who participate in your intervention, whereas the control or comparison group does not participate in any of your activities. The group is a control group only if random assignment was done. Otherwise, the group is called comparison group (Miksic, 2022). Age range, gender proportion and social criteria should be equally distributed in both groups.

The size of your sample (= intervention group + control or comparison group) depends on the population size of your intervention, the margin of error (confidence interval), the confidence level, and the standard deviation. This sounds like complicated statistics but don't worry, you won't have to calculate them yourself. You can find many sample size calculators → *online*, or you can use a statistical software.

Margin of error (confidence interval)

The margin of error is a percentage that indicates the extent to which you can assume that your survey results reflect the views of the population as a whole. The smaller the margin of error, the closer you are to an accurate answer at a given confidence level.

Confidence level

The confidence level indicates how certain one can be that the population would choose an answer within a certain range. For example, a confidence level of 95% means that there is 95% certainty that the results will be between the numbers x and y. The standard is 95%, but it can also be set at 99% or 90%.

Children & youth regularly take part in S4D activities →

Children & youth have not taken part in any S4D activities →

Inter-vention

Comparison/Control

Outcomes for both groups are measured and compared

How to measure outcomes

Requirements	Intervention group	Comparison/Control group
Regular participation in S4D activities	✓	x
Trainer/teacher is trained in S4D	✓	x
Older than 10 years	✓	✓
Mix of boys and girls	✓	✓
Mix of disabled and non-disabled children	✓	✓
Mix of social backgrounds	✓	✓
Mix of rural/urban context	✓	✓
Stable group	✓	✓

Requirements for the Intervention group and the comparison group

But what qualifies as an intervention in the sport sector?

- Programs with a set number of sessions over a fixed period provide a good structure for data collection (Petry, 2022).
- The total intervention length should at least be three months with regular weekly sessions. The number of sessions per week can be adjusted according to the length of the intervention, and you can plan more weekly sessions if the overall length is shorter (see table below).

Sessions per week	Months	Sessions in total
3 (1.5 h)	3	36
2 (1.5 h)	4,5	36
1 (1.5 h)	9	36

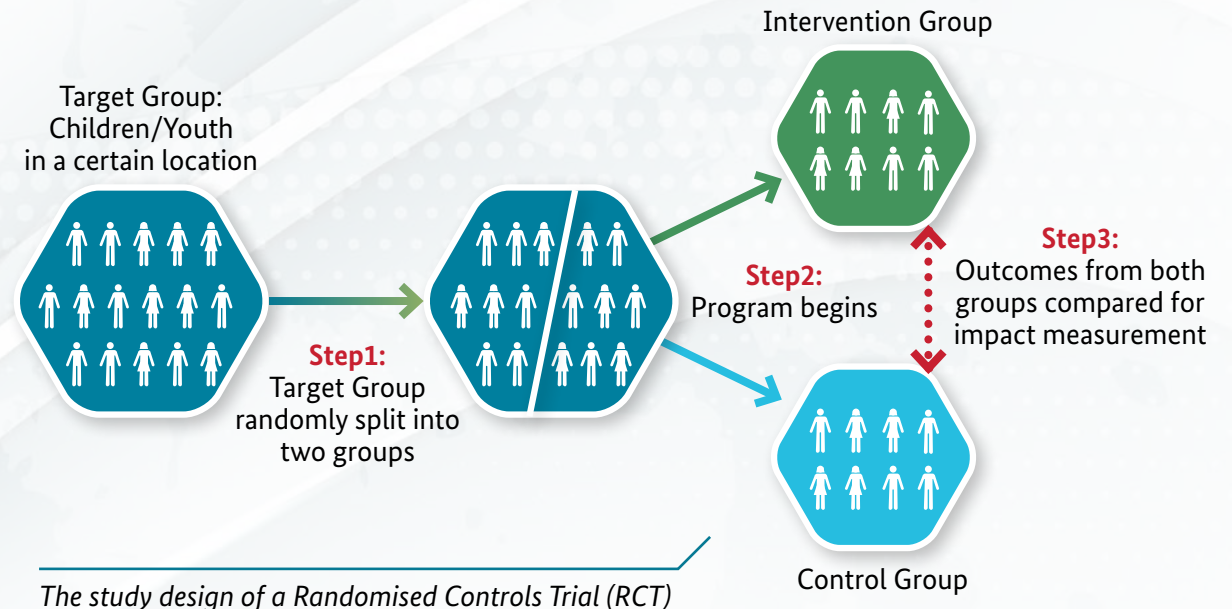*Ideal length and frequency of S4D Session*

The quality of the intervention depends heavily on the S4D coaches who carry out the training sessions

• Coaches must ensure that children/youth regularly attend trainings
  > *track attendance*
• The coach must formulate sporting and non-sporting learning objectives
  (life skills/competences) for every training session
• Life skills/competences are part of every training unit
  > *use planning and reviewing sheets*
  > *personal review after every session: consequences for the next training?*
• The reflection guarantees the transfer to the daily life
  > *use of different reflection methods (e.g., role play, in pairs, homework)*
  > *without a well-leaded, precise and in-depth reflection, there is no transfer*
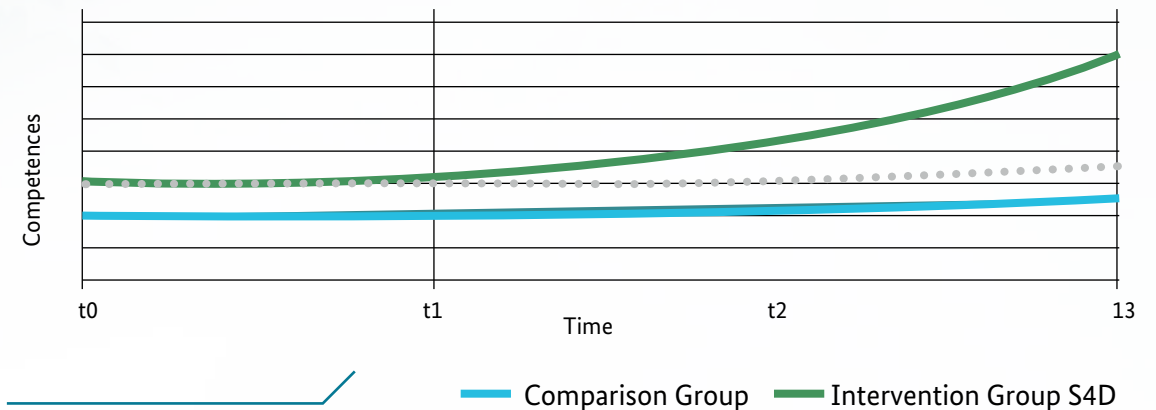
Study design

The study design aims to isolate the relationship between cause and effect (impact) from potential alternative explanations. The gold standard is the *experiment*, also known as **Randomised Controls Trial (RCT)**. The RCT is considered to be the standard trial for evaluating the impact of interventions since it provides the most reliable evidence. In an RCT, participants are randomly assigned into two groups of equal size by the researcher: the intervention and the control group.

Well-designed RCTs provide a powerful response to questions of causality since the findings can show that what is being achieved is a result of the intervention and not anything else.

Target Group:
Children/Youth
in a certain location

Step1:
Target Group
randomly split into
two groups

Step2:
Program begins

Intervention Group

Step3:
Outcomes from both
groups compared for
impact measurement

Control Group

The study design of a Randomised Controls Trial (RCT)

Both groups participate during multiple times *in the same survey* (at least twice: before and after the intervention):

Competences

t0   t1   t2   13

Time

Structure of an S4D survey

━━━ Comparison Group ━━━ Intervention Group S4D

In the end, you need to look at it from a practical angle and ask yourself what the realities in your project are and how you can implement a scientifically sound study in this setting.

The quasi-experimental design

The quasi-experimental design is a special form of an experiment. Working with this approach, the survey is conducted with existing groups, (e.g., existing training groups, classes, etc.) so there is no randomization. In a quasi-experiment design, the group that doesn't participate in the intervention is not called the control group, but rather the comparison group. In evaluation practice, quasi-experimental designs dominate. In evaluations that take place after or during an intervention, the intervention group has already been constituted by the project itself and cannot be randomly attributed. Also, there are cases of ethical boundaries that prevent deliberate random selection of group membership. For example, if one part of a community receives better access to water and another does not. Stockmann (2007) argues that as long as there are several points of measurement (before-after) and the conditions of the project implementation are known and run consistently, there are hardly any differences to the experiment in terms of design quality.

Challenges in setting up a Randomised Controls Trial or Quasi-Experiment

- *Access to a group,* where members *have similar characteristics* or traits (regarding social background, age, gender), that can be divided in two groups without harming either of them: intervention & control/comparison group.Ideally, the control/ comparison group is invited to participate in the program in the next year of activities.
- *Control/comparison group* members are not allowed to participate in any activity.
- Access to the group not only once, but at least twice (pre-and post-evaluation).
- Space and time to assess the groups in a *quiet environment.*

INPUT
Create information for the target group

> **Important: the participation is voluntarily – there are no disadvantages for not participating in the survey!**

Survey information

- Create a general survey information sheet for the target group with the following information:
  - ✓ Title of the study
  - ✓ Researcher/Research team in charge
  - ✓ Purpose of study
  - ✓ Procedures
  - ✓ Risks
  - ✓ Benefits
  - ✓ Confidentiality
  - ✓ Data protection guidelines
  - ✓ Compensation (if applicable)
  - ✓ Voluntary participation
  - ✓ Contact information
- Everyone who participates in the survey needs to be informed before the survey (if minors are involved, their legal guardians need to be contacted).
- This is compulsory for data protection!
- All schools or other 'gatekeepers' should be involved before the data collection.
- Depending on the data protection guidelines of each country or organization, there needs to be a signed declaration for participation (consent form), or in case of an online survey, agreement to participation via e-signature.
- In case of children and youth participating in the survey, the declaration for participation must be signed by the parents.

Literature

- Miksic, M. Y. (19. September 2022). *Johns Hopkins School of Education.* Glossary. Access *here*.
- Petry, K. (2022). *Impact Studies and the (ideal) length of S4D Interventions.* Internal Document
- Stockmann (2007): *Handbuch zur Evaluation. Eine praktische Handlungsanleitung.* Münster: Waxmann Verlag GmbH.

Resources to use

- → *Sample Size Calculator*
- → *The Easiest Online Form Builder (jotform.com)*

ASSESSMENT

SELF-LEARNING

1. How long should the total duration of the intervention be?

A : One month with regular weekly sessions
B: Three months with regular weekly sessions
C: Nine months with regular monthly sessions

2. What groups should your study design contain?

A: Two intervention groups
B: A control and a comparison group
C: An intervention group and a control or comparison group

3. What information does not belong on your information sheet for you target group?

A: Your targeted results
B: Purpose of study
C: Contact information

Task 1: Create an information sheet for your target group

Your information sheet should conduct the following information:

✓ Title of the study
✓ Researcher/Research team in charge
✓ Purpose of study
✓ Procedures
✓ Risks
✓ Benefits
✓ Confidentiality
✓ Compensation
✓ Voluntary participation
✓ Contact information

Task 2: Set up a consent form for your survey

SPORT FOR
DEVELOPMENT

Learning objectives

By the end of this step, you will be able to …

✓ … test your instrument before your data collection phase starts.
✓ … create a data collection plan.
✓ … understand of the survey conditions.
✓ … handle different data-collection-situations.

Step 6

Collect data

SPORT FOR
DEVELOPMENT

INPUT
Pilot testing and revision of quantitative instruments

Pilot Testing your data collection method

Before you start collecting data, you should always test your instruments – especially your survey – with a small group (max. 30 persons). Ideally, they have the same characteristics as your target group, but are not members of your target group.

Participants in the pilot cannot participate in the real survey! Talk to them about difficulties, problems, and their perceptions about the instrument – by doing so you will find out where adjustments of your survey are needed. You should also conduct a descriptive analysis of frequency distribution of your results to examine if there are any outliers or atypical response behaviour (e.g. everyone gave the same answer to the same question).

During your pilot testing you should ask yourself the following questions

- Are the questions clear and precise?
- Are your tools manageable to administer?
- Does the choice of words match the target group's own communication style?
- Have you checked if there is any content or if there are any questions that could make your participants feel uncomfortable or offended?
- Are the instructions for those carrying out the survey clear and unambiguous?
- Is the length and scope of the survey or questionnaire appropriate? Questionnaires and interviews that are too long can make respondents impatient and unfocused, which in turn has a negative effect on the answers' utility. This especially applies for children.
   > Less is more at the beginning.
   > Start with small, clear measures that you can later expand.

INPUT
Data Collection Process

Start planning your data collection at the same time as you are planning your programme!

Planning the data collection process

Survey conditions

- Trained interviewers ask standardised questions via a tablet/ smartphone (alternatively on paper) or self-assessment by the participants.
- The same children & youth.
- The same questions.
- Different points in time.
- Keeping everything else stable, the difference between the points in time can be seen as the impact.

Create a timetable

What?	Who?	When?
Define indicators and targets (develop frameworks SDG4+SDG5)		Dec 20
Select sample: Intervention and comparison/ control group		Dec 21
Develop quantitative questionnaire		Jan 21
Create survey information for target group		Jan 21
Organise data protection requirements		Jan 21
Translation of questionnaires		Jan & Feb 21
Programming of online survey		Feb 21
Select interviewers		Feb 21
Design interviewer training (e.g., hand-out, small presentation)		Mar & Aug 21
Train interviewers		Mar & Aug 21
Data collection: pre-test (quant)		Mar 21
Data collection: post-test (quant)		Aug 21
Data cleansing (quant. Interviews)		Sep 21
Quantitative data analysis		Sep 21
Design framework for FGD (based on quantitative results)		Sep 21
Data collection: focus group discussion		Oct 21
Data transcription (qualitative Interviews)		Oct 21
FGD data analysis		Oct 21
Data triangulation of quantitative and qualitative results		Nov 21
Write report		Nov 21
Publish results		Nov 21

Example of a data-collection timetable (Dec 20 – Nov 21)

Avoid collecting too much data and focus on the data you really need! Lots of data also means a lot of work!

Create a data-collection plan

> The **data-collection plan** helps you to structure and plan your data collection process. Remember: it is a dynamic process and needs to be adapted frequently.

How to manage challenges during the data collection phase

• The verification of long-term results is often challenging. It is hard to contact participants after much time has passed, and because of the many possible outside influences, it is difficult to verify that the results can be linked back to the programme.

  > Keep your participants' contact data up-to-date and inform/remind them that there will be a follow-up survey.

  > You can ask participants for their opinion on how the project has affected their current situation.

• Sometimes participants don't want to or are unable to take part in your survey. They might have concerns about participating, fear consequences, or not be able to provide information.

  > It is important to guarantee anonymity to all participants.

  > Perhaps you can interview people who have a close relationship with the participant (e.g., parents).

• When your programme targets organizations rather than individuals, the outcomes reflect the benefits that the organizations derive from the support provided by your programme.

  > The outcomes can be represented in part by quantitative data but qualitative data is most useful.

  > Medium-term outcomes are changes in an organization's knowledge, whereas increased effectiveness of the organisation's operations is a long-term outcome.

Resources to use

→ *Here* you can download you data collection plan template:
  Essentials – M&E in the field of S4D
• Phineo: → *Social-Impact-Navigator | Online tutorial for nonprofits*

ASSESSMENT

Task: Create your timetable

• Use the template:

What?	Who?	When?

SELF-LEARNING

Plan your Data Collection Process by setting up your individual Data Collection Plan.

• Use the link in the resources to download the data-collection-plan template.

Learning objectives

By the end of this step, you will be able to …

✓ … organize your collected data.
✓ … analyse your collected data.
✓ … derive recommendations for action.
✓ … know what to do in the event of unsatisfactory results.

Step 7

Analyse data & write reports

INPUT
Evaluation

INPUT
Data-Evaluation-Plan

Questions to ask yourself during your evaluation process

- Can you identify trends or clusters in the data?
- Were quality controls carried out during the processing of the data?
- Do the results differ so widely from the estimates that the processing may have been flawed?

- How do the results compare to other results?
- Are the results as expected, better, or worse?
- Are there any developments that differ from the intention of the programme? If so, why is this the case?

- Are there any identifiable connections?
- Are there any apparent changes in the basic assumption or context of the programme in relation to the needs of the target group?

- Does the project or the programme plan need to be adjusted?
- What should be done?
- What additional information is needed to answer the questions?
- Which points should be investigated or analysed in more detail?

> Analyse and evaluate your data as soon as possible, otherwise you run the risk of losing data, having out-dated data, or realizing too late that additional data is needed.

1. Organize your collected data

- Quantitative data can be prepared in tables (e.g., Excel, SPSS)
  > Answers should be coded by numbers (e.g., gender can be coded as follows: 1 is 'male', 2 is 'female', 3 is 'diverse' and 99 is 'no answer was given')

> The data from the participants must remain anonymous. Instead of names, use anonymized numbers or codes to organize the data.

Personal_Code	Age	Grade	Ppl_house	Frequency	GSE1_Score	GSE2_Score	GSE3_Score	GSE4_Score	GSE5_Score
ACGM19112004	15	7	8	0	1	1	2	2	2
AECM28072008	11	5	5	0	4	3	1	2	4
AFMA02082005	14	7	2	0	3	4	3	2	2
AJP05092008	11	5	7	0	2	3	2	3	2
AMRA17102005	14	6	0	0	3	2	4	2	1
ANLC04032007	12	5	7	0	3	1	2	2	2
ARC05052006	13	4	5	3	2	4	4	3	2
ASC18112007	12	5	10	0	3	3	2	2	3
ASRM14032008	11	5	6	0	2	3	2	4	2
ASSB13112004	15	6	3	3	4	2	3	4	2
AVA13112006	13	5	9	2	2	4	4	3	2
BADC15122005	14	7	3	1	3	4	2	2	3
BAGH07072007	12	6	6	3	2	4	4	3	2
BAGV08092002	17	6	6	1	1	1	2	1	2
BERS18082006	13	5	0	0	3	2	2	3	2
BLCI18112007	12	5	6	0	3	2	2	3	1
BSQO06102004	15	7	8	2	4	4	2	3	4

An example of quantitative data organised in an Excel table adapted from Soto (2020).

- For qualitative data (open questions): core statements are extracted and summarized (content analysis)
  > Transcribe audio data via software, e.g. *otter* or *happyscribe*.

2. Data analysis

- **Quantitative analysis:** descriptive statistics (averages/means, frequencies)
  - When analysing quantitative data, you must first decide which statistical software you are going to use. There is a range of software, such as SPSS, Stata, and R, and even Excel can be used for certain analyses.
  - Before starting with the analysis, you need to clean your data. Check for incorrect, corrupted, irrelevant, duplicate, or incomplete data and missing values within your dataset.
  - There are two types of statistics: descriptive and inferential statistics. Descriptive statistics describe the characteristics of a data set such as averages/means, frequencies, etc. Inferential statistics allow you to test a hypothesis and assess whether your results are generalizable to the broader population.
  - There is a wide range of statistical tests you can use to test your hypotheses. Remember, less is more! You do not need to conduct the whole variety of statistical tests for your study. Ask yourself what type of data you have and what you would like to analyse. For example, for a study design with two groups and at least two times of measurement, you could use the two-factor (two-way) repeated measures ANOVA.

- When analysing your data, you need to check for statistical significance but also for effect size. Testing for statistical significance tells us if an effect observed in the sample arose only by chance or sampling error and does not exist in the population (null hypothesis). A result is statistically significant when the p-value is below 0,05; in this case, you can reject the null hypothesis and assume that the effect did not occur by chance, with a residual probability for coincidence of 5% or less (depending on the p-value). The effect size determines if the effect is large enough to be meaningful in the real world, thus depicting the strength of association between the independent and dependent variable. As statistical significance is influenced by the sample size (a larger sample size increases statistical significance), you could obtain statistically significant results but with a small effect size. It is therefore essential that you also analyse the effect size. There are several measures for effect size such as Cohen's d, Pearson's r, and others.
- It is also important to interpret and explain the results of your statistical tests. Most of your readers won't understand the meaning of p=0.05.

> **Meaningful results need to be statistically significant AND show a large effect size.**

• Qualitative content analysis

Table

- A table can be used to organise data from questionnaires or Focus Group Discussions. The first column lists all questions from the interview, questionnaire, or FGD.
- In the first step all answers to the individual questions are summarized in the column 'Summary of answers'.
- In the next step recommendations are derived from the summary.

Results: Focus Group Discussion with Youth		
Questions	Summary of answers	Recommendations
O1: What do you **like** during your trainings?		
Q2: What do you **don't you** like during your trainings?		
Q3: …		

An example of a table from a Focus Group Discussion with Youth held in Albania

Thematic Analysis

- Thematic Analysis "is a method of qualitative analysis [widely used across the social and health science, and beyond] for exploring, interpreting and reporting relevant patterns of meaning across a dataset. It utilises codes and coding to develop themes" (Braun & Clarke 2022, p.224).
- Database: qualitative data e.g.: interview transcripts, focus group discussions, etc.

6 Steps to do a thematic analysis (Braun & Clarke 2022)

1 Familiarising yourself with the dataset: Get familiar with your dataset by listening to audio recordings, and reading & re-reading transcripts, interview and observation notes, and programme documents. Make notes about analytic ideas that appear to you.

2 Coding: Now work systematically through your entire dataset, always having your research question in mind. Identify codes of interview transcripts and observation notes by labelling certain relevant sentences or phrases to describe the content. After that, organise the relevant data referring to your codes.

3 Generating initial themes: Identify connections in form of shared core ideas and merge those codes. Construct broader themes for codes that address your research questions. You can use visual mapping to organise themes and codes.

4 Developing and reviewing themes: Go back to your dataset and review your themes in regard to the dataset. Check if your themes and codes make sense. Do they highlight the important information from the dataset? Reconsider your themes. Themes may be split up, collapsed or even be discarded. If you use visual mapping, redraw your thematic map.

5 Refining, defining and naming themes: Write down theme definitions and (re)formulate what exactly is meant by each theme.

6 Writing up: Write down your analysis of the data. Your text should contain an introduction section, a methodology section, a results section, and a conclusion section. Tell your reader the story about your dataset to answer your research questions.

Categorisation

• Categorisation is the most complex qualitative analysis. It consists of the following seven steps to be accomplished during the process of analysis.

1 initiating text work

2 development of main thematic categories

3 coding according to deductive categories

4 collection of text sections within the coded deductive categories

5 inductive determination of sub-categories

6 coding of the entire empirical data according to final set of categories

7 simple and complex analysis, visualisation

research question

Applied schema for a structuring content analysis (Kuckartz 2016, p.100)

• In the Coding template, dimensions are divided into codes. The codes can be divided into or assigned to sub-codes. They can also be described by code memos. In the last column an example quote can be placed.

dimension	code	sub-code	code memo	example
...	...	...	...	...

Coding template for a structuring content analysis

3. Define the results and draw overall conclusions and check your indicators and overall objectives

• Result presentation/illustration
  > Comparison, interpretation, and discussion (indicators and targets)
  > Statements of achievement or unfulfillment of indicators
• *Types of comparison:*
    1. Before-after comparison
    2. Target-actual comparison
    3. Comparison between various project configurations
    4. Comparison between groups
    5. Comparison among projects and benchmarking

4. Document your results and show your impact!

• Write a well-structured report
• Content of your report:
 • *Introduction*
 ✓ Background and context
 ✓ Explain your programme
 ✓ Explain the aim of the evaluation
 ✓ Explain your methodology
 ✓ Explain your sampling design
 • *Findings*
 ✓ Explain your findings
 ✓ Analyse and discuss your findings
 ✓ Show your impact
 ✓ Lessons learnt & process review (what went well and what did not?)
 • *Conclusion*
 • *Recommendations*

5. Communication of findings and derive recommendations for action

• Findings should be *widely* communicated and *discussed* in the context of a workshop (e.g., with stakeholders, community, target group, experts).
• Findings are channelled into *recommendations* that have direct impact on programme activities/which are used to adapt the programme.
  > Your stakeholders should identify with the recommendations to secure their support and the process of implementation.

What to do in the event of unsatisfactory results?

- Examine your evaluation process for inaccuracies.
- Systematic monitoring during the project should prevent you from being surprised by bad results.
- If that doesn't help, place the result in the project's overall context *use the problem tree (Step1)*.
- If your results still are unsatisfactory, it is time for transparency:
  > Communicate the results with your stakeholders.
  > Explain why the results are not as expected.
  > Plan countermeasures to improve the project and discuss them with your stakeholders.

Literature

- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide.* Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE Publications.
- Kuckartz, U. (2016). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (3., überarbeitete Auflage). Grundlagentexte Methoden. Weinheim, Basel: Beltz Juventa. Retrieved from → *here*
- Phineo (o.J.). *Social Impact Navigator. The practical guide for organizations targeting better results.*
- Calzolari Soto, V. (2020). *The Impact of the Golazo Programme from the Carvajal Foundation in Buenaventura/Colombia: A qualitative and quantitative study* (Unpublished master's thesis). Cologne, Deutsche Sporthochschule Köln.

Resources to use

→ *Qualitative Content Analysis*
- Example of an impact evaluation report: → *Galz & Goals Impact Evaluation Report Namibia 2016*
- Transcription software examples: → *otter* or → *happyscribe*

ASSESSMENT

True or false?

1. Quantitative data can be prepared in tables.

   True ◯   **False** ◯

2. Inferential statistics describe the characteristics of a data set such as averages/means, frequencies, etc. Descriptive statistics allow you to test a hypothesis and assess whether your results are generalizable to the broader population.

   True ◯   **False** ◯

3. Brown and Clarke defined 6 steps to conduct a Thematic Analysis: Familiarising yourself with the dataset, Coding, Generating initial themes, Developing and reviewing themes, Refining, defining and naming themes, and Writing up.

   True ◯   **False** ◯

4. There are different types of comparison to organize your collected data.

   True ◯   **False** ◯

5. Your report should include the following chapters: Introduction, Personal Opinion, Findings, Conclusion, and Recommendations.

   True ◯   **False** ◯

6. In the event of unsatisfactory results, you should manipulate your results.

   True ◯   **False** ◯

SELF-LEARNING

Task: Follow the Data Evaluation Plan step by step depending on your methodical approach and check your progress

To-Do	Check
1. Organize your data collected	
2. Analyse your data	
3. Define the results and draw overall conclusions and Check your indicators and overall objectives	
4. Document your results and show your impact	
5. Communicate your findings and derive recommendations for action	

KEEP ON TRACK

Commonwealth Secretariat: → *Sport and SDGs Indicator Framework*

> Category 3: Programmatic indicators – a common methodology for categorising programmatic or project-level interventions that produce results explicitly linked to prioritised SDGs and targets.
> The Category 3 indicators are not meant to act as a common approach to measurement, given the variety of programme types, methodologies, and contexts. Instead, we aim to utilise them to establish a common language to describe the type and depth of impact of sport, physical education, and physical activity on sustainable development.

Category 3: Programmatic Indicators

8 programmatic **impact areas** (Kazan Action Plan):
1. Sport for health and well-being for all [SDG 3]
2. Sport for making cities, inclusive, safe, resilient and sustainable [SDG 11]
3. Sport for quality education and lifelong learning for all [SDG 4]
4. Sport for peaceful, inclusive and equitable societies [SDG 10, 16]
5. Sport for economic growth and full, productive employment [SDG 8]
6. Sport for gender equality and empowerment for all women and girls [SDG 5]
7. Sport for sustainable consumption and action against climate change [SDG 12, 13]
8. Sport for effective, accountable and inclusive institutions [SDG 16]

ANSWERS ASSESSMENT (Step 1)

1. What is an impact?

A: An intended change only
B: A change of state, behaviour, or structure due to an intervention ✓
C: Negative and positive changes

2. What is an impact evaluation?

A: An analysis of outputs
B: Impacts, focal problem, causes
C: An analysis of medium and long-term effects ✓

3. Which three elements does a problem tree contain?

A: Effects, focal problem, context
B: Impacts, focal problem, causes
C: Effects, focal problem, causes ✓

ANSWERS ASSESSMENT (Step 3)

Name four different evaluation tools: (possible answers)

- Questionnaires
- Interviews
- Participant Observation
- Focus Group Discussion
- Story telling

- Diaries
- Blogs
- Photo or video monitoring
- Performing arts
- Painting and drawings

Where is the difference between a qualitative and a quantitative approach? Fill out the table:

	Quantitative	Qualitative
Procedure	Deductive, 'to measure': Counting, collecting numbers (e.g., demographic data)	Inductive, 'to understand the meaning': open questions, open tasks
Product	Hard, replicable data	Soft, realistic data
Process	Static	Dynamic
Perspective	External perspective of the researcher	Internal perspective of the people concerned
Interest	Explanation of causal relationships, generalisations	Exploration of the living environment, interactions, and relations

Name two validated instruments to conduct a standardized questionnaire:

1. _Self-Efficacy Scale
2. _PYD Questionnaire

ANSWERS ASSESSMENT (Step 4)

True or false?

1. It is beneficial to the study to use interviewers who are already familiar with the participants in order to build trust.

    True ⚪ **False** ⊗

2. Gender and age of the interviewers can influence the interview situation.

    True ⊗ **False** ⚪

3. Social desirability is the tendency to answer questions the opposite way a respondent feels.

    True ⚪ **False** ⊗

4. An interviewer training should not take longer than 3 hours.

    True ⚪ **False** ⊗

5. When using interviewers, a code of conduct is needed.

    True ⊗ **False** ⚪

6. Data protection is something that needs to be discussed with interviewers.

    True ⊗ **False** ⚪

ANSWERS ASSESSMENT (Step 5)

1. How long should the total duration of the intervention be?

A : One month with regular weekly sessions
B: Three months with regular weekly sessions ✓
C: Nine months with regular monthly sessions

2. What groups should your study design contain?

A: Two intervention groups
B: A control and a comparison group
C: An intervention group and a control or comparison group ✓

3. What information does not belong on your information sheet for you target group?

A: Your targeted results ✓
B: Purpose of study
C: Contact information

ANSWERS ASSESSMENT (Step 7)

True or false?

1. Quantitative data can be prepared in tables.

   True ⊗   **False** ◯

2. Inferential statistics describe the characteristics of a data set such as averages/means, frequencies, etc. Descriptive statistics allow you to test a hypothesis and assess whether your results are generalizable to the broader population.

   True ◯   **False** ⊗

3. Brown and Clarke defined 6 steps to conduct a Thematic Analysis: Familiarising yourself with the dataset, Coding, Generating initial themes, Developing and reviewing themes, Refining, defining and naming themes, and Writing up.

   True ⊗   **False** ◯

4. There are different types of comparison to organize your collected data.

   True ◯   **False** ⊗

5. Your report should include the following chapters: Introduction, Personal Opinion, Findings, Conclusion, and Recommendations.

   True ◯   **False** ⊗

6. In the event of unsatisfactory results, you should manipulate your results.

   True ◯   **False** ⊗

SPORT FOR DEVELOPMENT

Imprint

This guideline is based on the "Sport for Development (S4D)" web-seminars run by the
→ *German Sport University Cologne* in 2021/2022 as part of the cooperation with GIZ.
For further information on S4D check the → *S4D Resource Toolkit Website*.

SUSTAINABLE DEVELOPMENT GOALS

SPORT FOR DEVELOPMENT

german cooperation
DEUTSCHE ZUSAMMENARBEIT

gíz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Deutsche Sporthochschule Köln
German Sport University Cologne