



First Aid  
Reference  
Centre



# Outcome Measurement Toolkit

A toolkit for National Societies to support improvement of first aid education

Developed by the Global First Aid Reference Centre in collaboration with the British Red Cross

March 2021

Authors: Christine Boase, Ellen Gordon, Emily Oliver

*Outcomes Measurement can seem overwhelming and complex. We would encourage you to give it a go in whatever way you are able. There is always insight to be gained, even from projects that don't work out the way you planned. Our hope is for the toolkit to be a useful guide, providing a broad overview of outcome measurement. If you can't apply all of the toolkit, don't let this put you off, use the bits which work in your context.*

*If you have examples of measurement that you have applied in your National Society, please share them with GFARC so they can be added to this toolkit in future editions.*

## Contents

<b>What is the toolkit?</b> .....	2
How to use this toolkit .....	2
<b>Section 1: Principles of outcome measurement</b> .....	3
What is outcome measurement? .....	3
What might we <i>want</i> to measure? .....	3
What <i>can</i> we measure? .....	3
Definitions .....	4
Why measure learning outcomes? .....	5
<b>Wider value</b> .....	6
Who to measure .....	6
When to measure outcomes .....	6
<b>Section 2: Examples of outcome measurement tools</b> .....	8
Confidence Measures .....	9
Likelihood to Act Measures .....	10
Collecting data from learners .....	12
<b>Collection Methods</b> .....	12
<b>Section 3: Analysing the data</b> .....	14
Statistics .....	14
Comparisons .....	14
Questions to explore .....	14
Top Box Method .....	14
Average Potential .....	15
Using the insight .....	15
Brookfield’s Lenses .....	16
<b>Section 4: Case Studies</b> .....	17
Georgia Red Cross Society: pilot project to implement an outcomes measurement system .....	17
Testing a product – British Red Cross .....	20
Insights for data collection methods – Lithuania Red Cross .....	21
<b>Appendix 1: Theories of Outcome Measurement</b> .....	23
Kirkpatrick .....	23
Behaviour change models .....	23
Theory of change .....	24
<b>Appendix 2: Hypothesis Testing</b> .....	25
<b>Hypothesis Testing</b> .....	25
<b>Research questions</b> .....	25
<b>Appendix 3: Project planning – Example from the Andorran Red Cross</b> .....	26
<b>Contact and Global First Aid Reference Centre</b> .....	27

## What is the toolkit?

This toolkit gives an overview of approaches that national or regional first aid education teams can use to measure learning outcomes, and how the data can be used to improve the effectiveness of first aid education activity. The document focuses on the learning outcomes of lay-public learners such as children and adults in families, communities or workplaces. Many of the evaluation principles, tools and approaches are also adaptable to staff, volunteers or medical professionals who learn first aid.

It is intended that this document is a 'work in progress' and that it should be updated and added to as National Societies develop and test new and better ways of measuring outcomes.

## How to use this toolkit

- **Section 1: Background:** what are outcomes measures and how can we use them?
- **Section 2: Examples of measurement tools:** This section has several example measures which can be used and adapted as required to create outcomes measures for your own context.
- **Section 3: Analysis:** Read this section for an explanation of how to use the data to learn about and improve your education offer.
- **Section 4: Case studies:** This section details the experiences of National Societies which have trialled outcomes measurements. Read this to gain an understanding of how they applied the measures and the impact this has had.
- **Appendix 1: Theory and background.** Read this section for a deeper understanding of why outcomes should be measured, the benefits and challenges of application, plus an overview of some theories underpinning outcomes measurement.
- **Appendix 2: Further Analysis:** This section gives further guidance on how to use a more academic approach to data analysis.
- **Appendix 3: Project planning example** from the Andorran Red Cross

## Section 1: Principles of outcome measurement

### What is outcome measurement?

The core aim of education outcome measurement is to enable us to understand the change we make through the first aid education we provide. If we identify which outcomes we are most likely to affect with our education, and we measure how they change, then we can use that data as an indication of how effective our education is. Three categories of outcomes which could be used to indicate the effectiveness of first aid education are learning outcomes, patient outcomes, and societal outcomes.



#### Learning outcomes

- The changes that occur in learner's knowledge, skills, attitudes or beliefs as a result of their learning experience.



#### Patient outcomes

- The changes that occur in the medical condition of ill or injured people due to the first aid help our learners provide in an emergency.

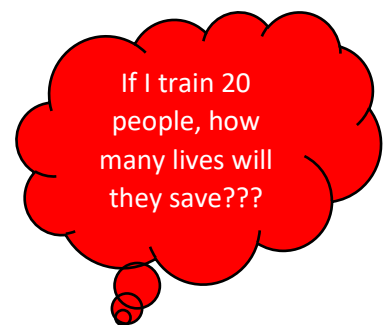


#### Societal outcomes

- The changes that occur in communities' resilience as a result of their learning experiences.

### What might we want to measure?

In an ideal future-world we will understand whether learners use their first aid skills to help others; how well learners perform first aid in an emergency; and whether ill or injured people suffer less, heal more quickly or live longer because of first aid interventions by our learners. However, the nature of first aid makes it challenging to measure the impact on patients and communities as they require tracking and monitoring learners over long periods of time or in-depth research. As such they provide very little evaluative feedback that first aid education teams can use to improve education effectiveness.



### What can we measure?

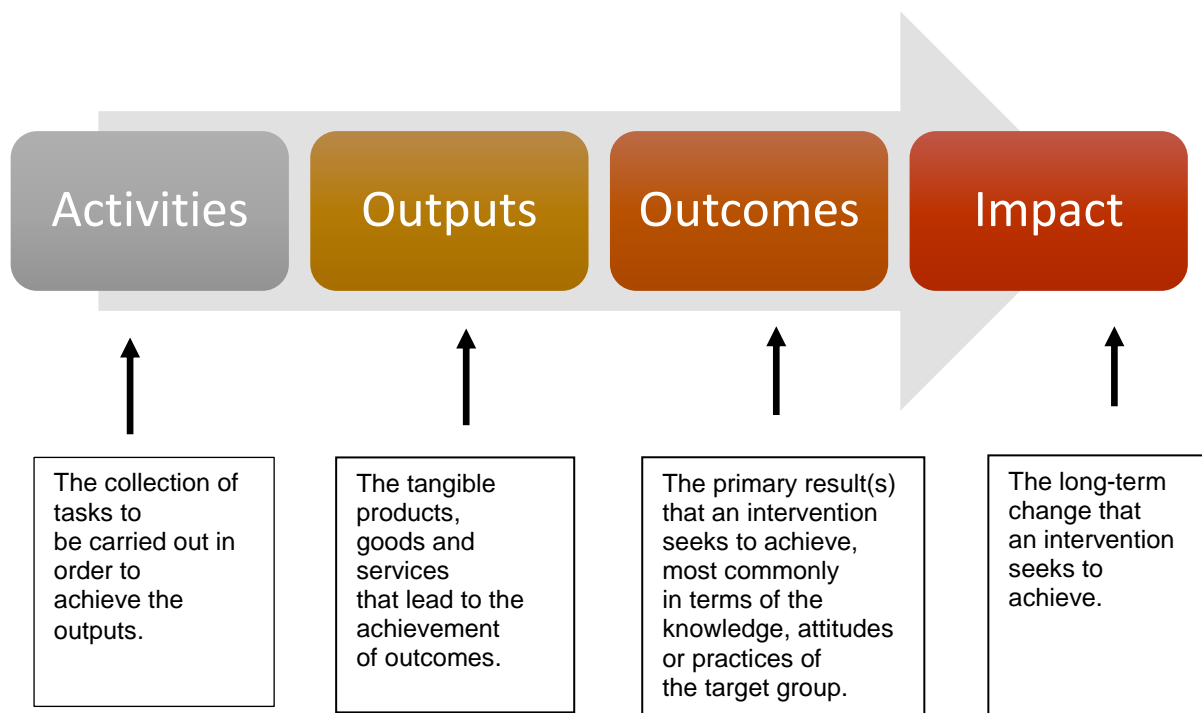
There is however one stage when we do have the ability to easily measure outcomes – and that is the learning experience; whether it is a face-to-face course, a self-directed digital course, or a blend of both. Learning outcomes such as knowledge, skills, attitude to helping can all be measured. For this reason, **this toolkit only focuses on learning outcomes**. This does not imply

- Knowledge
- Skills
- Attitude

that patient and societal outcomes are less important, just that they are harder to apply for organisational improvement.

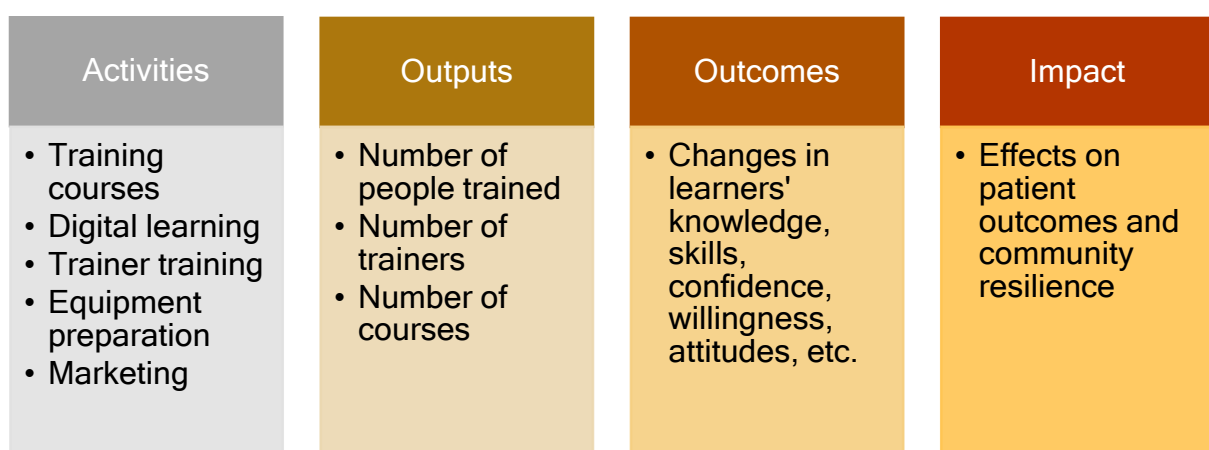
**Note:** ‘Attitude to helping’ is a term to capture the confidence or willingness of the person to act, or the likelihood that they will act. It is ambiguous because it is dependent on the situation at the time of the emergency and is therefore hard to predict in advance. Nonetheless, we have found that it is an important measurement because we can use it to determine if the education they have received has changed their attitude. To measure a change in attitude, you should articulate what you mean by this in a way that is relevant to your culture.

## Definitions



*Results chain with definitions. Reproduced courtesy of IFRC Project/Programme planning - Guidance manual*

Building on the IFRC programme and planning results chain (above) we can define the first aid education programme results chain as follows.



*First aid education results chain*

Many National Societies already report on their outputs (i.e. things you can count) such as the number of people trained and number of active first aid education volunteers. However, while outputs can tell us about the *size* of operational activity, they do not give any indication of the *quality* of the education. To understand the quality of Red Cross / Red Crescent education, we must evaluate outcomes.

Is this trainer improving after the coaching I have given them?

### **Why measure learning outcomes?**

When choosing which first aid course to go on, people (consumers) will consider the quality of the first aid education offer. To date, the Red Cross has relied upon brand reputation to equate with quality, however this is becoming increasingly difficult in competitive markets. To retain our leadership in first aid we must make efforts to monitor and evaluate education quality.

It is also only through evaluating learning outcomes that we can really start to improve the effectiveness of first aid education, so that learners are more likely to respond in a first aid emergency. It can be resource intensive to set up and maintain the ongoing evaluation of learning outcomes. Investment is required in the processing system, the ongoing reporting and the subsequent improvement of programmes. But if our aim is to create people and communities who are more able, confident and willing to respond to first aid emergencies, then this is an investment worth making.

Consider what specific benefits your organisation wants to see from measuring outcomes and have those clear in mind as you set up an evaluation system. The value of measuring learning outcomes may be described through the lens of the value it could bring to learners, to the organisation, or the wider value for the Movement.

#### **Learner value**

Well-structured evaluation ensures that learners understand what they going to learn in the session ahead. It enables them to engage with the topic and make connections to the learning outcomes anticipated for them. These aspects combined may increase their motivation to learn. At the conclusion, evaluation then provides time for learners to reflect on what they have learned and whether they have achieved their outcomes. This may contribute to the consolidation of their learning, and the identification of further learning they may feel they need. Learners may also feel some satisfaction at having had the opportunity to feedback on their experience.

I understand what I am going to learn.

#### **Organisational value**

The benefits of introducing a system which evaluates learning outcomes are perhaps greatest when viewed from an organisational level. Evaluation data can give performance trends over time if recorded consistently or a snapshot of performance at a moment in time if done periodically. A thoughtfully implemented evaluation system can also provide outcomes data that allows analysis of performance through variables such as learner demographics, individual or cohorts of trainers, course length or type, or region. At an organisational level, measuring learning outcomes data can be used:

- > to identify and monitor evidence-based improvement projects
- > to help develop educators' teaching practice
- > to help evaluate the effectiveness of new or existing products
- > to provide valuable evidence to funders, consumers and decision makers
- > global cross-border comparative research studies, and collaboration in curriculum design
- > to carry out research projects to determine for example which activities or methodologies are most effective.

Is this course effective for this audience?

### Wider value

Imagine a world where all first aid education providers evaluate learning outcomes; where huge amounts of data is used to understand how to make our practice more effective. In the global world we live in, much like the fields of communicable diseases or climate change, it is only through mass global participation in evaluating outcomes that organisations can set benchmarks and targets to improve. The Red Cross is perfectly placed to lead the way in doing this. Ultimately, leading on quality will help secure our brand reputation, raise the standards of first aid among all other first aid education providers and thereby ensure more lives are saved, and suffering is reduced.

### Who to measure

There is value in measuring the outcomes of any person who is learning something whether it is a member of the lay public or a long-serving volunteer. However, the learning outcomes we choose to measure and the processes we use to measure them may vary between learner audiences according to their context, needs and motivations for learning. The following sections focus on measuring the learning outcomes of lay-public learners (such as children or adults in families, public or workplaces).

The number of learners you measure is very much down to the capacity you have to process the evaluation data. The more learners you can measure, the greater and more reliable the data insight you are going to have. An ideal would be to aspire to one-day evaluating all learners, all the time. However, if you have limited resources, consider consistently evaluating a manageable proportion of learners, evaluating everyone but only at set intervals each year, or carrying out specific research projects to understand the quality of your programmes.

### When to measure outcomes

To measure a learning outcome, it is recommended that the evaluation questions are asked before the learning takes place, and then again after the learning experience has finished. The before (pre) and after (post) evaluation enables us to be able to compare the two results and the change that has occurred. If you are only able to ask the questions after the learning experience, it is important to capture the

- Ask the same questions *before and after* the learning takes place.
- Compare the results to understand the change

learner's perception of change.

There is no definitive right or wrong in terms of how long before or after the learning the evaluation surveys should take place. However, to be able to attribute the change in outcomes directly to the learning experience, evaluation should take place as near to the course as possible. There is evidence that suggests people's outcome values start to deteriorate within days of learning first aid. To understand the immediate outcomes of learning it is better to do the post evaluation as soon as possible after the learning experience.

On that note, it can also be very useful to understand the retention or maintenance of outcomes over time. So particularly where electronic systems are being used, you could consider setting up a further retention evaluation survey that is sent to learners 3 – 6 months after learning. This would ask the same evaluation questions to enable it to be compared to the original pre and post data.

## Summary

- Ensure learners are clear about what they are going to learn at the start
- Decide what learner outcomes you want to measure: knowledge, skills, attitude to helping (confidence, willingness, likelihood to act)
- Ask the same questions at the start and end of their learning experience so you can measure the change



## Section 2: Examples of outcome measurement tools

In this section we take two concepts of outcome measurements and show you different ways to use them. As most National Societies already have their own mechanisms in place for assessing the knowledge and skills of learners, we are going to focus on two less-used outcomes of **confidence** to act and **likelihood** to act.

- By **confidence** to act, we mean that the learner feels happy in themselves that they will act effectively in a first aid emergency.
- By **likelihood** to act, we mean that the learner feels it is likely that they will step forward to help in a first aid emergency.

To use these concepts to understand the difference that the education has made, we need to ask the learner before they learn and after they learn. We can then measure the change that the education has made.

These concepts can both be measured using a very general question such as this:

***How confident do you feel that you can act effectively in a first aid emergency?***

General questions give an overall measure of the concept of first aid. They can be a useful starting point for measuring general effectiveness. However, you need to remember that every single learner will have a different interpretation of this question, such as the seriousness of the emergency, the type of injuries they might have to deal with, and where it happens. This means that it will be more difficult to analyse and learn from the data you collect.

- ✓ General measure of effectiveness
- ✗ Every learner will interpret it differently

You can narrow these variables and make your data analysis tighter by using a vignette or scenario-based question such as this:

***You are on a quiet road and see a man fall off his bike. He has banged his head and is bleeding heavily. How likely are you to stop and help the man?***

This type of question places the learner in a much more specific scenario with less room for their own imagination to shape the situation. Answers help you to understand confidence levels regarding a specific skill (e.g. dealing with bleeding) and give insight into contextual considerations. You can use vignettes

which draw specifically on the context your learners might find themselves in, which can help you to tailor your education to deal with any particular concerns.

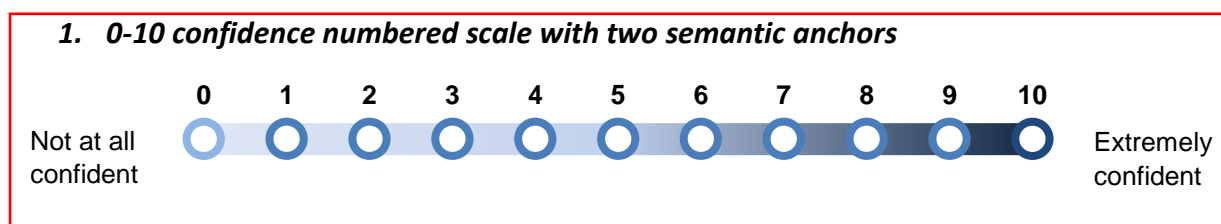
- ✓ Specific
- ✓ Broadly consistent interpretation
- ✗ More limited in what it tells you

## Confidence Measures

**General question:** *How confident do you feel that you could help effectively in a first aid emergency?*

**Scenario One Helping someone you love:** Your two-year old son is eating his lunch. While you are making him a drink, you notice he has gone quiet. When you look at him, he seems to be choking. How confident are you to help your son in this situation?

**Scenario Two Helping a stranger:** Someone in a cafe has just collapsed suddenly. You saw that he hit his head on the corner of a table as he fell. He is now lying very still on the floor and there is a small pool of blood forming from the wound on his head. How confident do you feel to help?

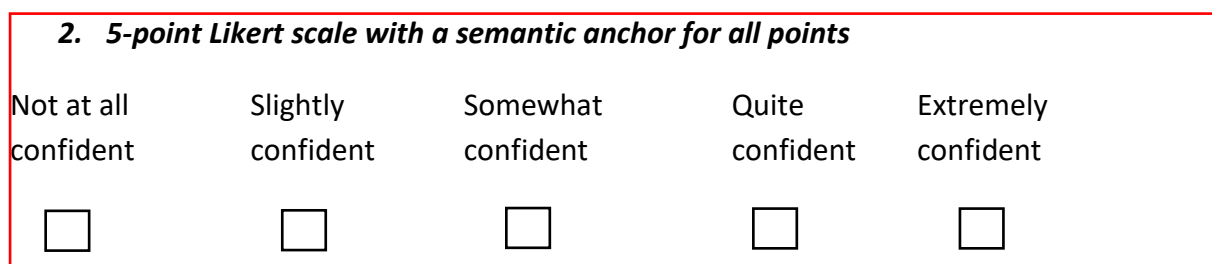


Advantages:

- > A 0-10 scale allows for greater differentiation in responses, so is often said to be a more sensitive scale than one with less response points.
- > People don't always read semantic anchors, so only having two (one at each end) increases the likelihood they will be read and understood.

Disadvantages:

- > Some people argue that a 0-10 scale is confusing and or overwhelming for respondents because it gives them too many options to choose from.
- > Not naming all the response points allows for greater individual interpretation of what each point represents.



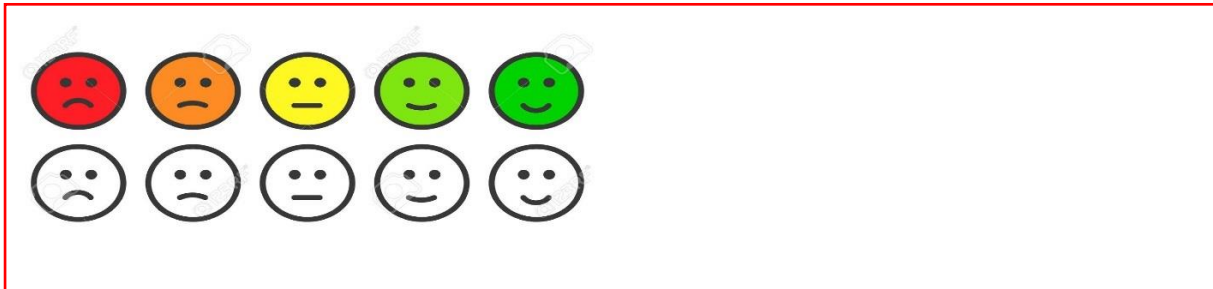
Advantages:

- > Respondent has less options to choose from meaning they are less likely to be overwhelmed or confused.
- > Naming each point means that there is less individual interpretation.

Disadvantages:

- > Some people argue that a 5-point scale does not allow for enough diversity in answers, this then leads to clustering (the majority of answers ending up at the same point) usually around the 3/4 points.

### **5-point Likert scale with visual anchors**



Advantages:

- > Helpful in a setting where literacy levels/language is a barrier

Disadvantages:

- > Lots of room for individual interpretation
- > Smiley faces could be interpreted to represent a construct such as happiness, rather than the one being measured, in this case confidence

### **Things to Consider**

- How clear and understandable the anchors are to your target audience
- Whether visual anchors will be more useful if the target audience is diverse, e.g. children and adults.
- Is confidence understood as a concept in the context of your target audience?

### **Likelihood to Act Measures**

**General question:** *How likely would you to be act in a situation that requires first aid?*

**Scenario One** *In a setting with other people present:* You are on a busy city street doing an important errand on your lunch break. An older woman walking by herself stumbles and falls heavily to the ground a few meters in front of you. How likely are you to stop and offer help?

**Scenario Two** *In a setting where no one else is present:* You are on a quiet road and see a man fall off his bike. He has a cut on his face and has hurt his arm. How likely are you to stop and offer to provide first aid assistance?

**1. 7-point numbered scale with two anchors**



Advantages:

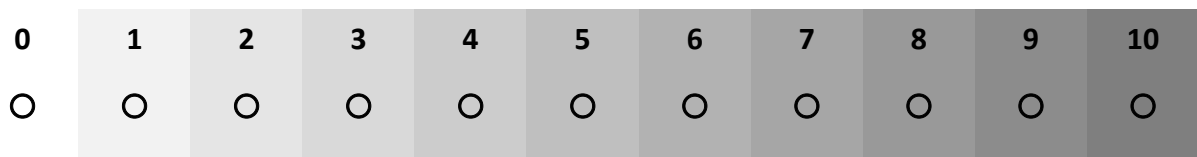
- > More sensitive (more options) than a 5-point scale, but not as many (not as overwhelming as a 0-10-point scale).

Disadvantages:

- > Same issue regarding not all points being given a semantic anchor.

**2. 0-10 Likert scale with three semantic anchors**

Extremely unlikely    Neither likely nor unlikely    Extremely likely



Advantages:

- > Mid-point has a semantic anchor, meaning it is more likely to be correctly interpreted.

Disadvantages:

- > Similar to the disadvantages on the 0-10 confidence scale.

**Things to Consider**

- Are there social factors in your scenario which might make someone more or less likely to act?
- Action could mean calling for help or giving first aid treatment, if you want to find out whether someone will take a specific action, this needs to be explicit in the question.

## Collecting data from learners

Administering evaluations can be time consuming and if done poorly, can lead to inaccurate or incomplete data. It is therefore vital to give educators guidance and support.



For educators to administer evaluations successfully, they need to understand what they are administering, why they are doing it, and how it should be done.

What:

- What is being measured?
- What is the construct and how does it relate to first aid?

Why:

- Why is this being measured?
- Why is it important to measure outcomes at all?
- What are the benefits?

How:

- How do I administer this evaluation?
- How do I make sure learners feel comfortable enough to answer honestly?
- How do I answer questions from learners about how to answer?
- How will the insight from the evaluation be used?

Educators should be given clear instructions on when evaluations should be used during the learning experience. This could even take the form of a script for them to follow. Guidance on not allowing personal bias to affect learner responses should also be included. A script to follow may go some way towards reducing any display of personal bias.

### Collection Methods

There are a range of ways that you can collect answers to the types of questions shown above.

- > **Post-it notes** Learners could be given a post-it before and after the learning to stick on the wall to represent their position on the scale.
- > **Physical lines** Learners could physically stand in a line on a scale (marked out with anchors by pieces of paper). The educator could then take a photo before and after the learning to mark which point each learner was at.



- > **Raising hands** The educator could read out in point on the scale in order. Learners would then raise their hand when the point they agree with is read out. The educator would then need to mark these responses down.



### Things to Consider

- The above methods of data collection are not anonymous, so other learners would be able to see individual responses. This might mean people are less likely to answer honestly.
- These methods require the educator to be organised both in terms of noting down the responses and making sure they can match an individual's before and after responses.

### > Forms

If you have the resources to use forms, here are some considerations.

	Advantages	Disadvantages
<b>Electronic evaluation</b>	Automated system Automated processing	Not every learner will complete the forms. High drop off rate of form completion after course and even higher for a retention study. Potentially costly to buy/develop electronic system. Potentially limited by tech
<b>Paper form evaluation</b>	Can ensure that every learner completes before and after. Occurs right before and after learning -consistent timing for all learners.	Potentially difficult and costly to get the paper forms returned and processed in a central place. Need to buy software to process the forms. Very difficult to study retention.

#### Batch headers

- These are identification codes which mean you can identify one set of evaluation forms from another (for example, courses held on different dates, different trainer etc).

#### Independent variables

- These are other pieces of information you might like to capture about your learners to help you compare results between different population groups e.g. age, gender, profession.

#### Anonymisation

- Answers should remain anonymous. Instead of using names, individual forms should be denoted by a unique code. This also applies to post-it notes. Some collection methods make anonymisation more difficult.

## Section 3: Analysing the data

The example tools on the previous page are all Likert scales. Likert scales are psychometric rating scales which are used to measure people's responses to questions. To analyse your results from them, it is important to understand key characteristics of data of this kind.

Scales which have an ordered range, such as smiley faces or descriptions are called 'ordinal'. Below are some approaches you could use to analyse the data you will get from this kind of scale.

### Statistics

It is not possible to calculate the mean (average) for ordinal data. If you don't know the distance between the categories, you won't be able to find the average. For example, it would be impossible to say that the distance between 'not at all confident' and 'somewhat confident' was the same as the distance between 'quite confident' and 'extremely confident'.

You *can* calculate:

- The mode – the most common response
- The median – the middle response when all responses are arranged in order
- The range/interquartile range – to show variance

You could also use a bar chart or a frequency test.

### Comparisons

If you are conducting the evaluation before and after the learning experience, this provides you with an opportunity to measure the change that has taken place. This could be as simple as looking at the percentage of learners whose confidence levels improved after the learning experience. Comparisons could also be made between the outcomes generated by different types of learning experiences or the outcomes achieved by different groups of learners.

Comparison could be done on an individual basis, or you could use the statistics described above and compare the most common response of the group (the mode) before learning and after learning.

### Questions to explore

You could try and analyse your data through answering some key questions such as:

- > How many people had a higher confidence level after the education?
- > How many people had the same confidence level before and after the education?
- > How many people had a lower confidence level after the education?

### Top Box Method

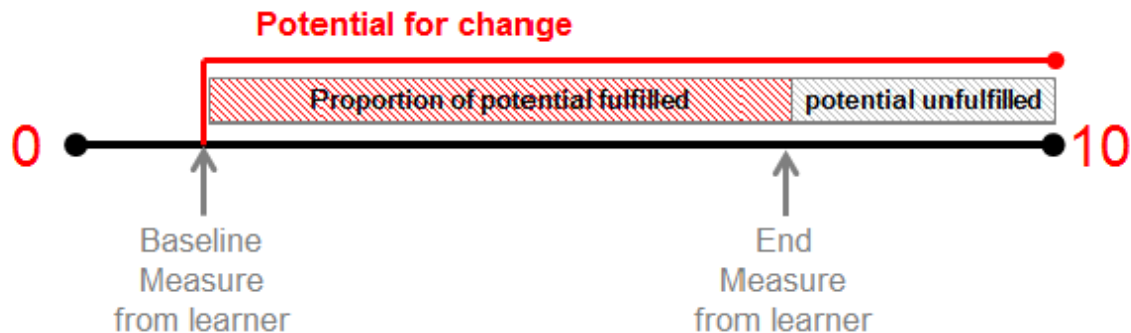
As explained above, it is not possible to calculate the average for this type of data. Instead, you could use the 'top box method'. In this approach, you look at what percentage of learners selected the top two scores on the scale (9 and 10). This is known as the top 'N' box where N refers to the number of scores you have included in your top box. If you include 9

and 10, this is a top 2 box. You could then look at the percentage of learners who scored in the top 2 box before the learning experience, compared to the percentage of learners who scored in the top 2 box after the learning experience.

### Average Potential

Average Potential is a validated tool developed by the British Red Cross. It can be used to measure any change pre and post a learning experience. It works like this:

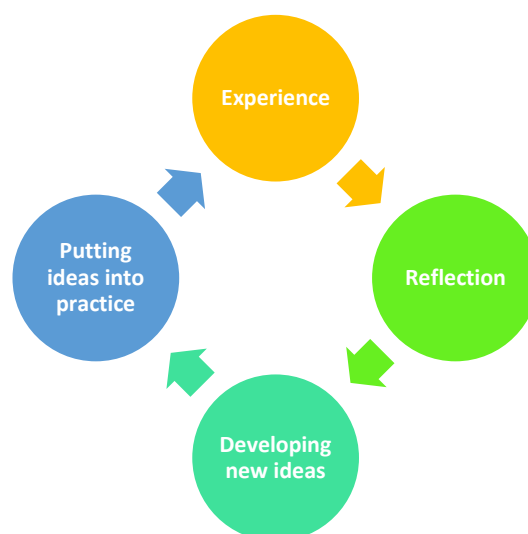
#### Calculating Educational Effectiveness



In the above example, the learner has rated themselves at point 2 on the scale (0-10). This means that their potential for change at the start is 80%. The learner then rates themselves at point 8 at the end of the course. Point 2 to point 8 is 60% of the scale. The education effectiveness of the intervention is 60/80 or 75%. This is the approach used by the British Red Cross to analyze the effectiveness of its first aid courses. The data is used to identify which trainers and which methodologies work best for which learner audiences. The National Society can then apply best practice across its courses to raise the standard of the education offered.

### Using the insight

Once you have analysed your data, you have a valuable tool for use in developing your education offer. The aim is to apply the insight as part of a reflective cycle, continually examining and adapting your offer based on the needs and context of your learners, and the outcomes you wish to achieve.



Kolb's Reflective Cycle



## Brookfield's Lenses

A common motivation for evaluating first aid education is to analyse and develop trainer performance. Outcomes measures can be very useful for this but work best when complemented by other forms of analysis. Brookfield's Lenses provides a useful model for how to do this. Analysing and developing trainer performance only through the lens of learner evaluation data can be demotivating for trainers. Brookfield suggests that for trainers to improve their practice that they must be able to analyse their own assumptions about their training through four different lenses (Brookfield, 2017).

### Brookfield's Lenses

**Self-reflection** – trainers can focus on their own experiences as learners or their own experiences as a trainer. In practice this could mean writing training journals, setting and reviewing personal training goals or reflecting on learner feedback.

**Student's eyes** – understanding the student experience is of 'utmost importance' to Brookfield. In practice this could mean facilitating learner' evaluations, doing focus groups with learners, and listening to feedback (verbal and written).

**Colleagues perceptions** – peers can highlight traits that have remained hidden despite using self-reflection and student feedback. In practice it might mean receiving observation coaching, talking with peers about practice, or doing some team-teaching.

**Theory and research** – Brookfield argues that a critically reflective teacher will draw on scholarly literature. In practice this could mean that trainers read, research, publish or present literature on training practice.

## Section 4: Case Studies

### Georgia Red Cross Society: pilot project to implement an outcomes measurement system

Georgia Red Cross Society (GRCS) carried out a project to contribute to the evidence base related to implementing an outcome evaluation system. The objectives were to set up an evaluation system, evaluate the results, and capture the key challenges and learning.

The project had three stages:

- 1) planning and development,
- 2) implementation,
- 3) analysis.

The following case study will summarise each of the stages including key activity, challenges, and any learning or findings.

#### Planning and Development

GRCS wanted to understand the quality of their first aid training in the different regions of the country. They identified *confidence and willingness to act* in an emergency as the outcomes they wanted to measure. As they were using paper evaluation forms, they also took the opportunity to ask learners about their likelihood to recommend the course (net-promotor score), the ratings of some aspects of the training experience, and for general feedback comments.

The batch header form (completed by the trainer) captures region, learner audience type, course duration, and the trainer ID. The evaluation form and batch header were created with the support of the British Red Cross who use the data capture software TeleForm. This software can design paper evaluation forms and capture the data from evaluation forms after learners have completed them. The evaluation form and batch header is available in English and Georgian. Forms were printed at the head office and posted to Branch offices.

The project was coordinated from the GRCS Head Office who communicated the aims and objectives with all branches. During an annual Training of Trainers gathering, trainers were given a learning bite on evaluation which explained how to introduce the evaluation forms to learners, and how to collect them back. This included a role play of using the evaluation forms in a course context. Trainers were encouraged to share any challenges they encountered during the project.

#### ➤ Challenges at this stage:

- Logistics including printing, postage and forms being returned for analysis.
- The scanning of the forms also proved challenging due to resource limitations.

# Evaluation form

The evaluation form, completed by each learner asked them the following:

- How confident do you feel that you can use these skills in a real emergency? 0-10 (Asked before AND after)
- How willing are you to help someone in a real emergency? 0-10 (Asked before AND after)
- How would you rate the training delivery for each of the following?
  - Presenting information clearly
  - Balance of theory to practical work
  - Providing constructive feedback
- How likely are you to recommend Georgia Red Cross first aid training to your friends and colleagues? 0 - 10
- Please tell us more about your answers. ((Free text))

FAEEN



power  
indness

## Implementation

The trainers were responsible for giving out the evaluation forms and ensuring that learners completed them before and after their course. The trainers then collected them back and returned them to the head office.

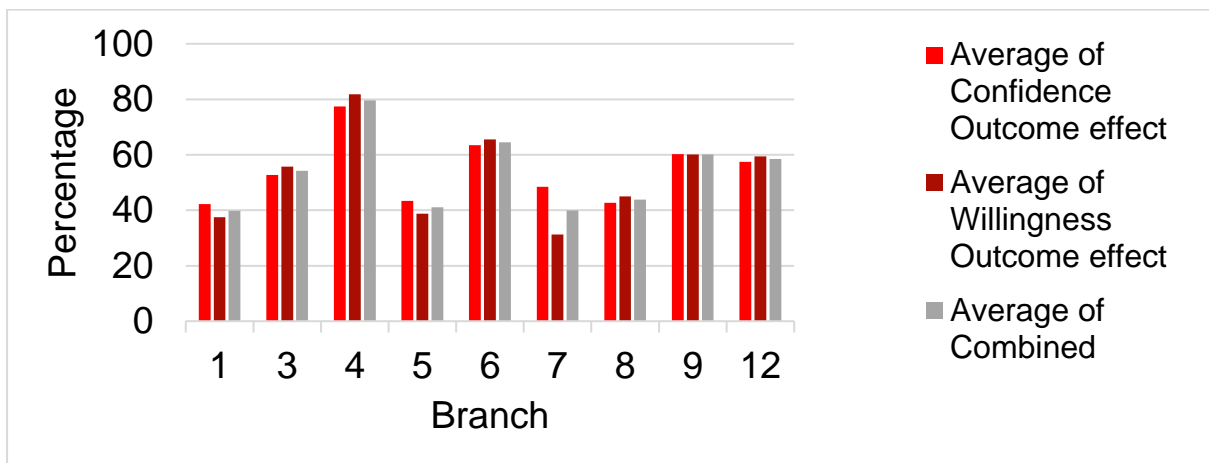
- Challenges at this stage:
  - An excess of paper to manage where branches were also using their own evaluation measures.
  - Ensuring participants filled out both the pre- and post-elements of the measure also proved challenging in some cases.

## Analysis

The forms were posted to the British Red Cross to be processed. 3200 forms were processed, data was analysed and then shared with GRCS. 'Average Potential' measurement was used as a mechanism for analysing the effectiveness.

The GRCS calculated confidence and willingness levels and then segmented the results to get an insight into differences between branches, learner type, course length amongst other variables. Below are some examples of their data analysis.

**Fig. 1 Confidence and Willingness Levels segmented by Branch**



The graph above shows the effect of different branches on confidence and willingness. The national average for confidence, willingness and for them both combined was 54%.

**Fig. 2 Confidence and Willingness Levels segmented by learner group**

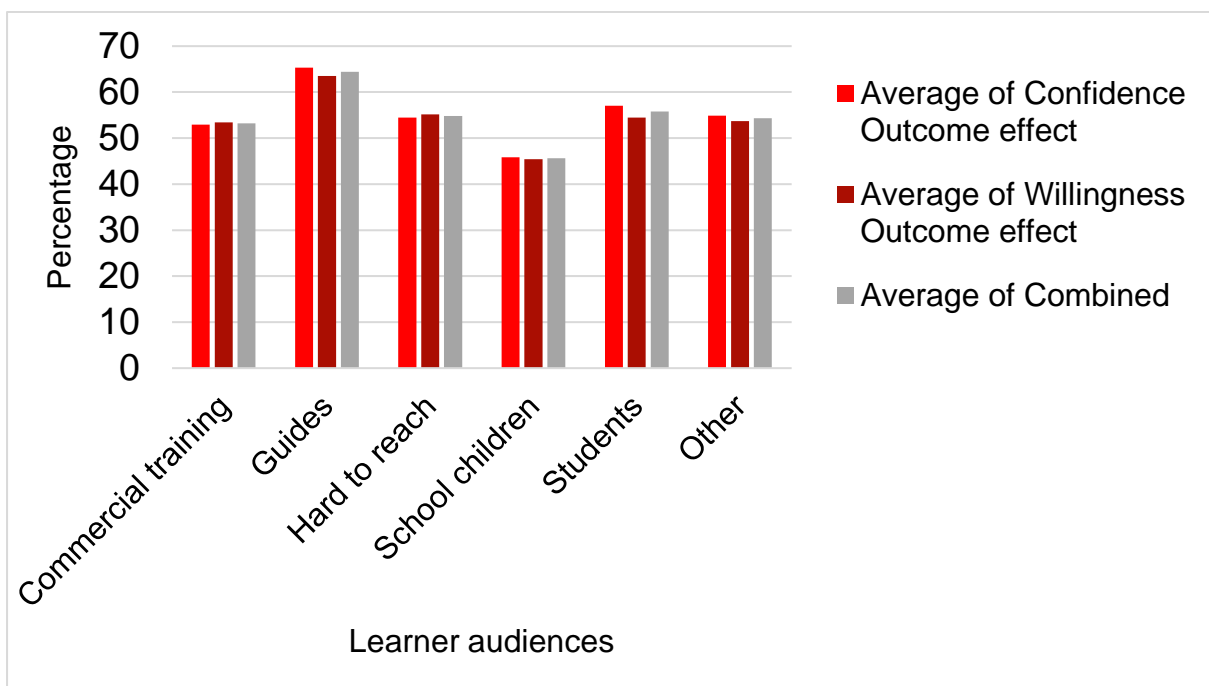
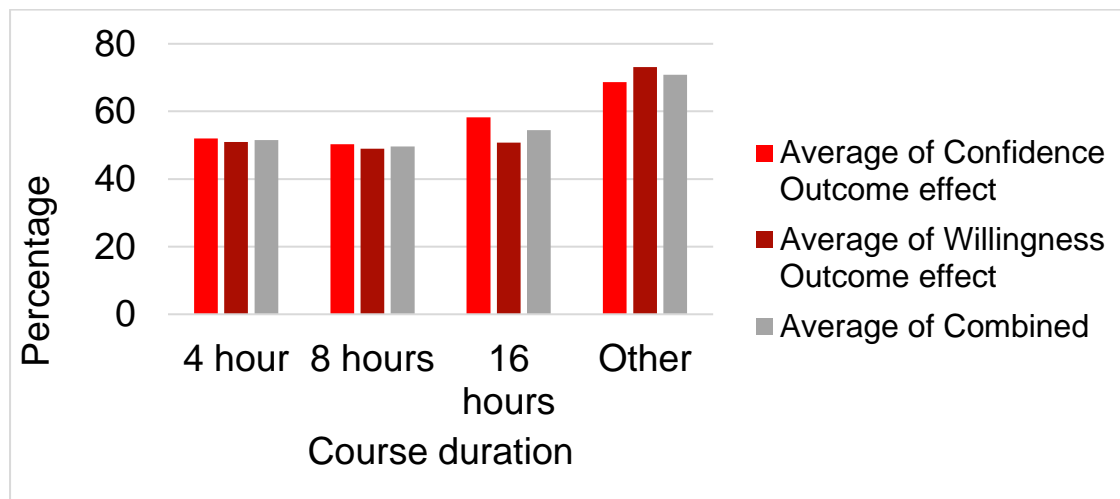


Fig. 3 Confidence and Willingness Levels for Different Course durations



### Application and Learning

The GRCS used the insight generated by this project and adapted part of their approach to make learners feel more confident to act in an emergency.

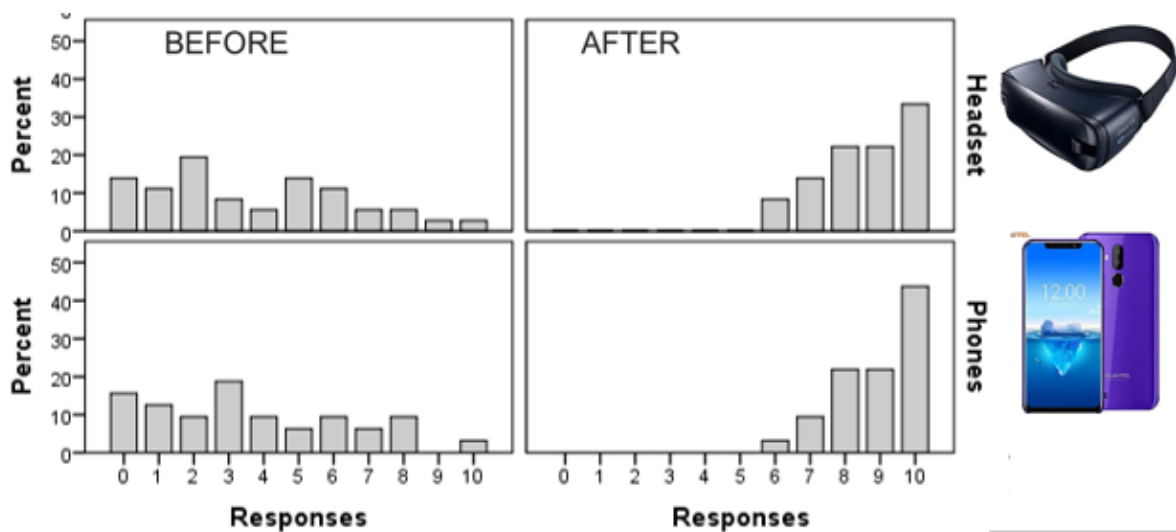
‘This showed us that we have to do more to focus on (people’s need to be more) confident. We changed our approach from, if you do something wrong you will kill the person, to, if you help you (could) save them’.

Whilst confidence proved to be a useful and relevant construct, in this context, willingness proved to be difficult to translate. Learners did not really understand the construct and the point was raised that culturally, this might not be the most appropriate thing to measure. This is crucial insight and should not be dismissed as a negative result.

### Testing a product – British Red Cross

At the British Red Cross we used outcomes measurement to help us make decisions about how we used our 360 bystander video within courses. First, we tested the effectiveness of the video as a tool to increase likelihood to act. Each learner was provided with a headset on which to watch the video and they self-rated their likelihood to act in a particular situation before and after the first aid course that they attended. Through analysis, we revealed that learners liked the video and that it was effective at changing minds about helping. But there was also a financial question. We wanted to know if we really needed to invest in headsets for it to work well, or if learners could just watch it on their phones. For this test, as shown below, the pre-learning distribution for likelihood to act is similar for both cohorts, and so is the post learning. If anything, the phone was marginally more popular than the headset. This was good news as it meant we did not need to invest in a lot of headsets for learners.

**You are walking through a busy shopping centre when you see someone you don't know collapse. How likely are you to give first aid?**



This very basic kind of testing can be used to test the effectiveness of different activities within a course (such as role play or no role play; or use of a quiz or no quiz).

### Insights for data collection methods – Lithuania Red Cross

The Lithuanian Red Cross attempted to measure learning outcomes by testing both paper and online surveys for learners. In developing the project they achieved an adequate level of buy-in from the educators who were responsible for ensuring that learners completed the surveys because they recognised that the insight provided would be beneficial for them.

In one year they collected about 600 before-and-after questionnaires on paper, as part of a training project for a corporate customer.

The challenge was that the paper forms were found to be difficult to process – the National Society did not have the scanning equipment and software available and therefore the scheme required a lot of data input which was time consuming and risked opportunities for mistakes to be made.

The National Society then moved to developing online questionnaires with a small sample. Pre-course surveys were sent out with registration information two days before the course, and follow up surveys were sent out two days after the course.

The surveys asked:

“How inclined would you be to help?”

“How confident would you be to help”?

“How likely, in your opinion, are your [coworkers/friends/etc.] to help you?”

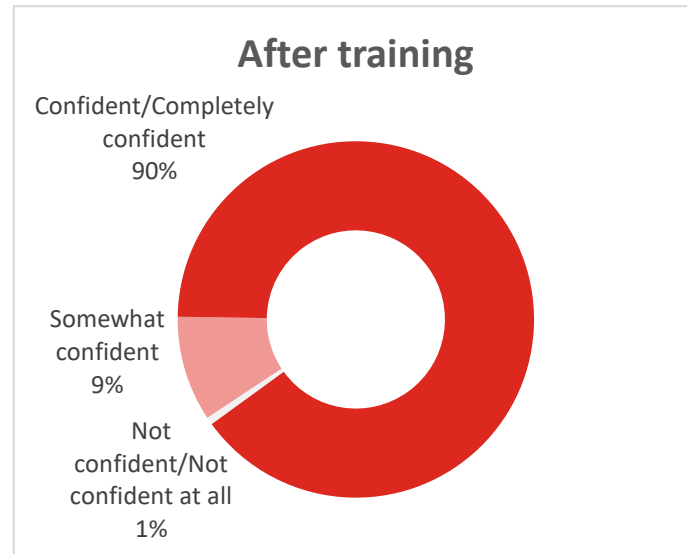
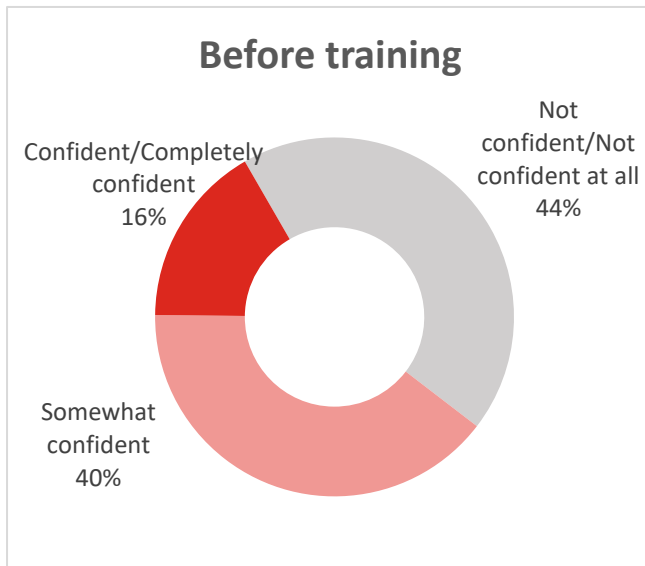
The rationale for these questions was that the National Society wanted to see the change in community relations and trust, as this was seen as a major part of training value.

All three questions are asked in a specific context of three situations - loss of consciousness, choking or injury. Follow up survey also includes net promoter score question and open-ended invitation to provide more details.

An example of how the data that was collected can be presented is below:

## How confident would you be in your skills at this time to help someone who is choking?

May-June 2019, N=576



	Before training	After training
<b>Not confident/Not confident at all</b>	252	4
<b>Somewhat confident</b>	229	55
<b>Confident/Completely confident</b>	95	517

## Appendix 1: Theories of Outcome Measurement

### Kirkpatrick

The Kirkpatrick is a widely used training evaluation model. It provides a useful structure for thinking of the things we can evaluate. It bears some similarities to the programme evaluation approach the IFRC uses.

#### The Kirkpatrick Model (Kirkpatrick D. , 2009)

Level 1: Reaction: The degree to which participants find the training favourable, engaging and relevant to their jobs

Level 2: Learning: The degree to which participants acquire the intended knowledge, skills, attitude, confidence and commitment based on their participation in the training

Level 3: Behaviour: The degree to which participants apply what they learned during training when they are back on the job

Level 4: Results: The degree to which targeted outcomes occur as a result of the training and the support and accountability package

4. Results

3. Behaviour

2. Learning

1. Reaction

Level  
one

*Figure 1 Kirkpatrick's Four levels of training evaluation. Reproduced from Kirkpatrickpartners.com*

(reaction) and level two (learning) can both be evaluated as part of the learning experience. Level three (behaviour) and level 4 (results) are more difficult to measure as they occur after and away from the learning experience. This toolkit will provide ideas on evaluating levels one and two.

### Behaviour change models

The IFRC use behaviour change theories to underpin the educational approach to first aid education (IFRC). The Theory of Reasoned Action (TRA), Theory of Planned Behaviour (TPB) and the Integrative Model of Behaviour Prediction (IMBP) are all theories of behaviour change. These models were developed sequentially, each building on the previous. The initial TRA was based on the idea that intention to act is the best predictor of behaviour (Ajzen & Fishbein, 1980). The TRA aimed to describe behaviours that were under a person's control (Sheppard, Hartwick, & Warshaw, 1988) based on the assumption that behaviour was voluntary. Behaviour is not always voluntary and so the additional construct of perceived behavioural control was added to the model, developing the model into the TPB (Sheppard, Hartwick, & Warshaw, 1988). An integral aspect of the TPB, like the TRA, is the idea that the stronger the intention of a controlled behaviour, the more likely a behaviour will happen

(Ajzen, 1991). The TPB is still one of the most frequently used and influential models for behaviour change (Ajzen, 2011).



The IMBP is a development of the TRA and TPB. It extends the previous models and includes the constructs of skills and environmental barriers as influencing intentions. It also expands upon the idea of normative beliefs and their role in determining behaviour (Yzer, 2012). The model accounts for both rational and irrational behaviours as it proposes that all behaviour is reasoned from the beliefs that person holds about said behaviour. The main concept of the IMBP in line with the previous model iterations, is that intention to perform behaviour is the biggest predictor of actual behaviour (Yzer, 2012). This intention is informed by beliefs about the behaviour, the individual’s attitude, perceived norms and self-efficacy.

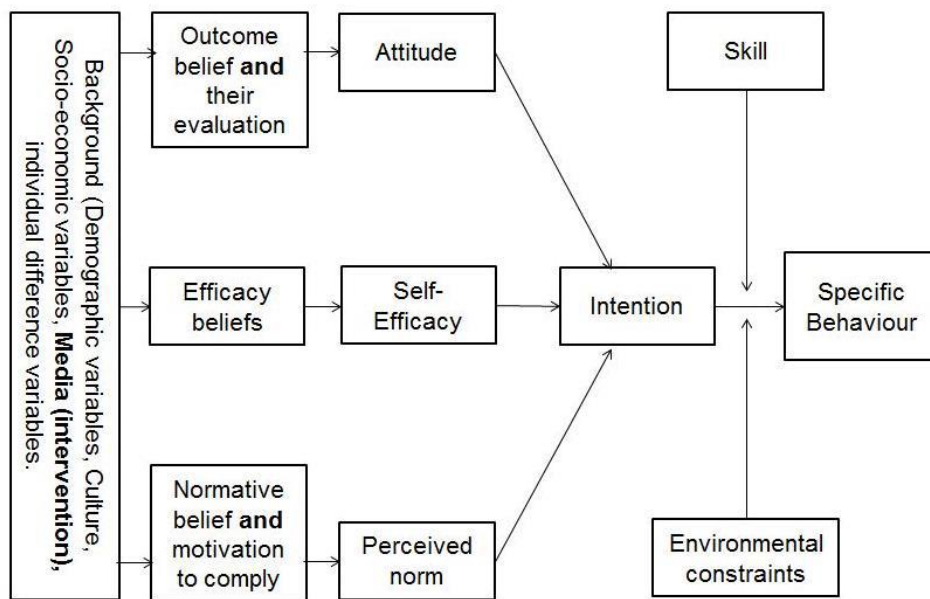
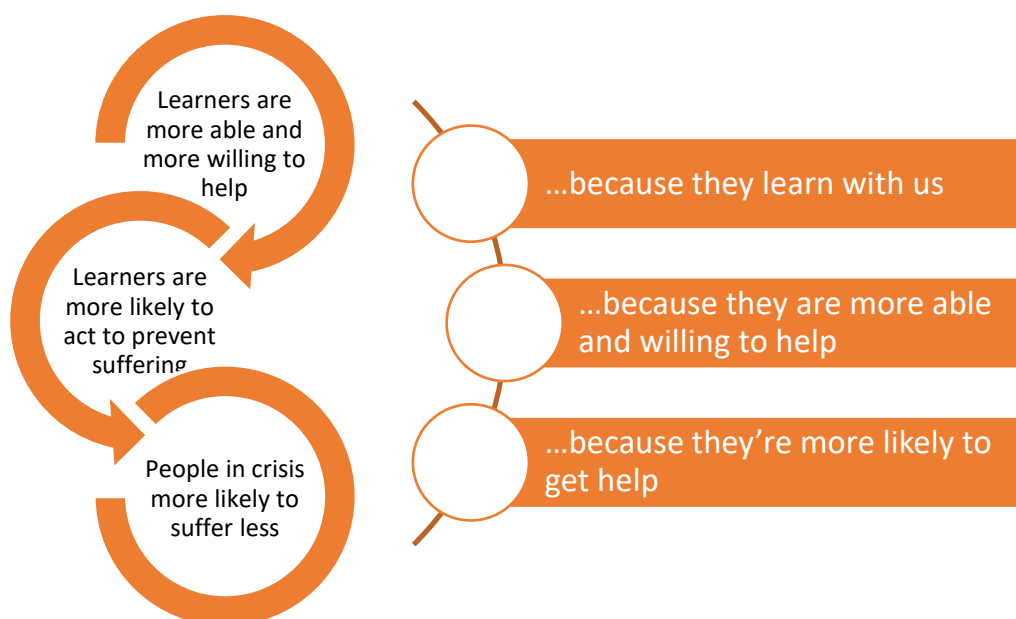


Figure 2 The integrated model of behaviour prediction

## Theory of change



## Appendix 2: Hypothesis Testing

### Hypothesis Testing

You might want to explore whether there is a statistically significant difference between two groups. One example of this would be comparing the confidence levels of two different learner groups. Analysis of this kind is usually carried out either through parametric or non-parametric tests. Within academia there is debate around which type of test should be used with scale data, although this is unlikely to be relevant to you unless you would like to publish your analysis. There is a growing consensus that it is acceptable to use either parametric or nonparametric tests as they have nearly equivalent type 1 error<sup>1</sup> rates. Some researchers would recommend using both kinds of test for the same data set.

The following tests can be calculated on programs such as Excel or SPSS.

- 2 sample t test (parametric)
- Mann Whitney U (nonparametric)
- Kruskal Wallis (nonparametric)

These tests will indicate whether you can accept or reject the null hypothesis. This is the hypothesis that there is no difference between groups. Rejecting it means there is a statistically significant difference. An example null hypothesis:

*There is no difference between the confidence levels of lay people who learn first aid through a phone app, and the confidence levels of lay people who learn first aid through a face to face course.*

### Research questions

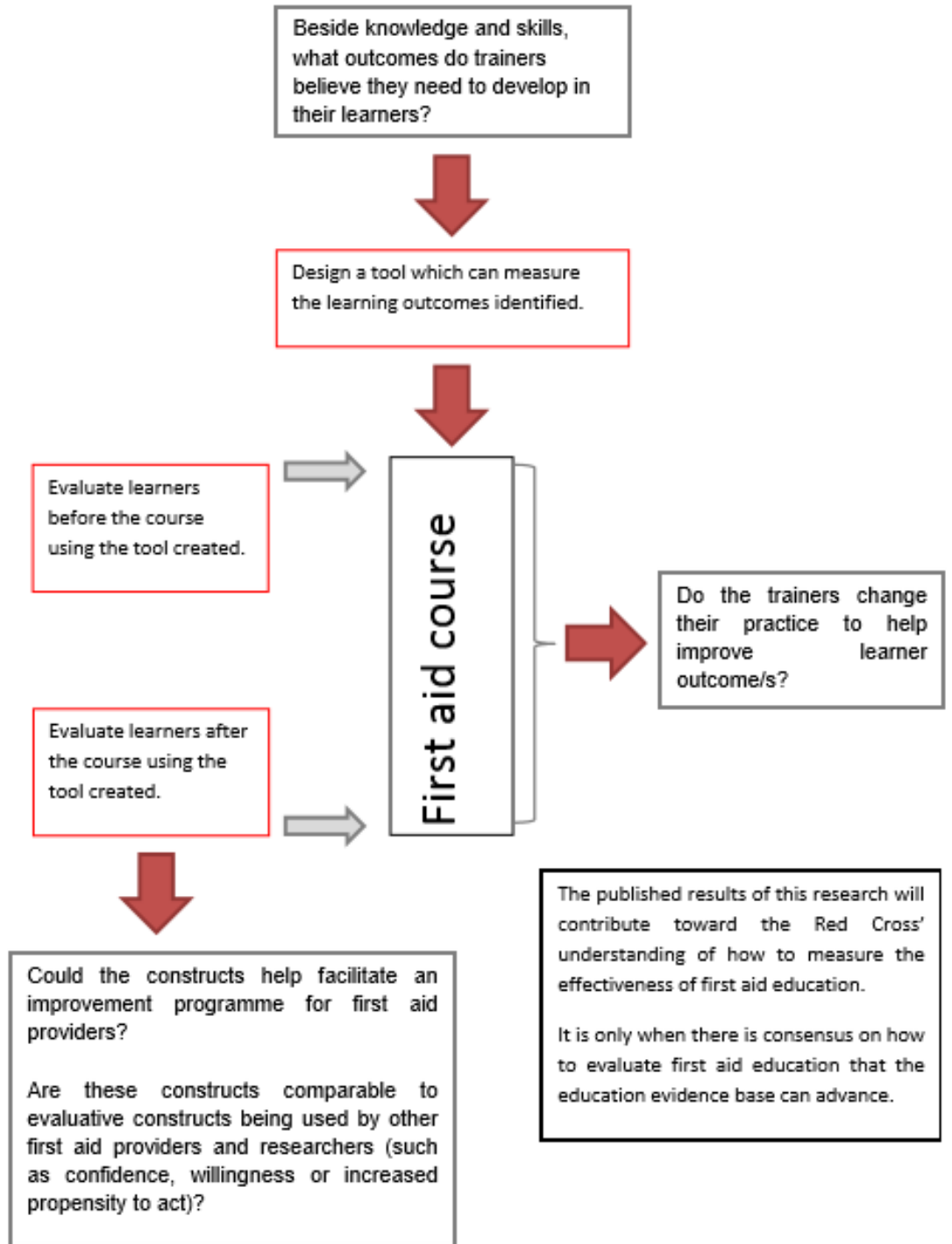
It can be helpful to set a research question at the start of a project on outcomes measures. This question can be used to guide analysis, including what to measure and what aspects of the data are relevant. Below are some example research questions.

- In this first aid course, is *Activity A* more or less effective at improving outcomes than *Activity B*?  
*e.g. In this first aid course, is role play more or less effective at improving outcomes than a group debate?*
- In this first aid course, is *Methodology A* more or less effective at improving outcomes than *Methodology B*?  
*e.g. Is a blended learning experience more or less effective at improving first aid outcomes than a face to face learning experience?*
- Does 'this' specific learner audience achieve better learning outcomes with A versus B course structure/methodology/activity?  
*e.g. Are parents first aid confidence levels improved through use of a first aid phone app?*
- Is this length of course more less effective than that length of course?

---

<sup>1</sup> A type 1 error is where a null hypothesis is rejected even though it is true, i.e. a statistically significant test result suggests a population effect exists when it does not. This is due to random sample error.

### Appendix 3: Project planning – Example from the Andorran Red Cross, French Red Cross (supported by Global First Aid Reference Centre & Fondation Croix-Rouge)



**Quiz used for this project :**

**Link learnings and interventions**

Age:

Training course date:

Name of the trainer:

Are you pursuing this training course because of a professional demand? Yes No

Have you already experienced/ or been a witness for a situation that required First Aid?

Yes No

Have you already been raised awareness of the importance of pursuing the First Aid Certificate ? Yes No

Capacity to learn						
This section's goal is to measure if First Aid gestures taught are easily assimilated by the learner and if he identified the most important gestures						
	Questions	Completely in disagreement [1]	In disagreement [2]	Neither in disagreement nor in agreement [3]	In agreement [4]	Completely in agreement [5]
1	First Aid gestures taught are easy to do					
2	First Aid gestures taught are easy to remember					
3	Fist Aid gestures taught are clear and without ambiguity					
4	I know important gestures I have to remember to be able to intervene					
Sharing responsibilities						
This section's goal is to measure if the learner is conscient of his legal obligation to provide assistance knowing that it allows him to not be the only responsible of the victim						
	Questions	Completely in disagreement [1]	In disagreement [2]	Neither in disagreement nor in agreement [3]	In agreement [4]	Completely in agreement [5]
5	I have the legal obligation to alert emergency services to aid a victim					
6	By calling emergency services, I get psychological support for my intervention					
7	By calling emergency services, I get technical support for my intervention					
8	By calling emergency services, the responsibility of the victim's treatment is shared (I don't feel I'm the only one responsible of the victim)					
IMPLEMENTATION						
This section's goal is to measure if the learner is conscious of the positive impact of his intervention on the victim's health						
	Questions	Completely in disagreement [1]	In disagreement [2]	Neither in disagreement nor in agreement [3]	In agreement [4]	Completely in agreement [5]
9	I have to intervene quickly as possible for the victim (if it is allowed by the situation)					
10	Each First Aid technics taught respond to a precise aim					
11	First Aid steps taught led to the most efficient treatment for the victim					
12	My intervention can save lives					

FINAL						
	Questions	Completely in disagreement [1]	In disagreement [2]	Neither in disagreement nor in agreement [3]	In agreement [4]	Completely in agreement [5]
13	As the outcome of the formation, are you feeling ready/ prone/ volunteer to provide assistance to a victim if a First Aid situation appears to you?					
14	As the outcome of the formation, I'm glad of the importance of my role as a citizen-workplace first-aider in my country					

**\*\*\*\*Questions asked can be similar. Nevertheless, you have the choice to answer them similarly or not. There is no obligation of coherence in your answers\*\*\*\***

## Contact

About the [Global First Aid Reference Centre](#): The IFRC Global First Aid Reference Centre aims to develop first aid training in accordance with the Movement's recommendations and inter-national scientific guidelines. The GFARC also focuses on supporting National Societies in delivering first aid training in their individual countries and facilitate network-wide information sharing, ensuring quality management of first aid and supporting first aid harmonization within the Movement.

Contact: [first.aid@ifrc.org](mailto:first.aid@ifrc.org)

Facebook: [Global First Aid Reference Centre](#)



**First Aid**  
Reference  
Centre