



INnovative and New Solutions for Sustainable Employability

INTENSE

EVALUATION VON BILDUNGSPROGRAMMEN
LERNINHALT

ARE YOU INTENSED?

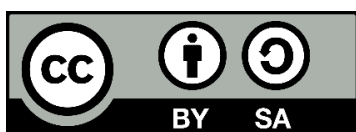


Co-funded by the
Erasmus+ Programme
of the European Union

Index

1	EINLEITUNG.....	2
1.1	EVALUATION EINES BILDUNGSPROGRAMMS.....	2
1.2	GRÜNDE FÜR EVALUATIONEN	2
1.3	EVALUATIONSPLAN.....	3
2	BASISKONZEPTE.....	7
2.1	CIPP MODEL.....	7
2.2	CONTRIBUTION ANALYSIS	8
2.3	OECD KRITERIEN	12
2.4	THEORY OF CHANGE.....	12
3	EVALUATIONSDESIGN.....	14
3.1	INTRO.....	14
3.2	KAUSALE ATTRIBUTION.....	15
4	INDIKATOREN.....	16
4.1	DEFINITION.....	16
4.2	CHARAKTERISTIKA VON INDIKATOREN	16
4.3	TYPEN VON INDIKATOREN	17
5	BERICHTE, KOMMUNIKATION UND ETHISCHE ANFORDERUNGEN	19
5.1	BERICHTE.....	19
5.2	KOMMUNIKATIONSSTRATEGIE	20
5.3	ETHISCHE ANFORDERUNGEN	21

Version 1.0 by the INTENSE consortium - veröffentlicht: 19-08-2019



Unless otherwise stated, this work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

1 EINLEITUNG

1.1 EVALUATION EINES BILDUNGSPROGRAMMS

Wenn wir über "Programmevaluierung" sprechen, beziehen wir uns auf eine Maßnahme, die aus zwei Teilen zusammengesetzt wird: die **EVALUIERUNG** und die **BILDUNGSPROGRAMME**.

Die Evaluation bedeutet, Urteile zu fällen oder Bewertungen vorzunehmen, um auf deren Basis bestimmte Entscheidungen zu treffen.

Ein Bildungsprogramm kann man als Zusammenstellung von personellen, methodischen, materiellen und zeitlichen Ressourcen beschreiben, für ein Bildungs- oder Kompetenzziel von Personen in einer konkreten sozialen Realität.

Programmevaluation kann daher verstanden werden als:

Geplante und systematische Maßnahmen, die auf ein zielgerichtetes Bildungsprogramm angewendet werden. Dabei wird die Absicht verfolgt, eine Wertbeurteilung für einen Teil oder für die gesamte Entwicklung, Planung oder Wirkung dieses Programms abzugeben und daraus nützliche Informationen für Entscheidungen zur Wirksamkeit, Fortführung oder ggf. Anpassung des Programms zu erhalten.

1.2 GRÜNDE FÜR EVALUATIONEN

Die Umsetzung eines Programms setzt die Beteiligung von Menschen voraus, investiert Zeit und Ressourcen und beeinflusst unser enges soziales Umfeld.

Ein Programm durchführen bedeutet also eine Investition zu tätigen, Ressourcen zu mobilisieren, die während der verbleibenden Zeit nicht für andere Aktivitäten verfügbar sind. Dabei ist es wichtig den langfristigen Einfluss des Programms sicherstellen und die Auswirkungen zu prognostizieren und natürlich laufend zu überprüfen.

Eine Evaluierung ermöglicht es und, den Nutzen des Programms sowohl für Personen (Beobachter, Schüler, Lehrer, Unterstützer usw.) in den beteiligten Institutionen als auch im Umfeld / Netzwerk zu ermitteln. Sie ist ebenso ein wertvolles Instrument, um die getätigten Investitionen zu begründen.

Zusammenfassend kann man die folgenden Gründe für die Durchführung einer Evaluation nennen:

- ✓ Mitteleffizienz öffentlicher Gelder
- ✓ Gesicherte Entscheidungsgrundlage
- ✓ Programmverbesserungen
- ✓ Überprüfen und Dokumentation kurzfristiger und langfristiger Auswirkungen
- ✓ nachhaltig Kompetenzentwicklung bei Klienten und Schülern
- ✓ Möglichkeiten zur Adaption auf andere Bildungszusammenhänge
- ✓ usw.

Mithilfe der Evaluation werden Aspekte identifiziert, die nicht auf dem erwarteten Niveau funktionieren. Daneben ermöglicht sie auch, die Probleme zu identifizieren und herauszustellen, die im Programm korrekt funktionieren. Schließlich können sogar die Problemstellungen identifiziert werden, die im eigentlichen Programm nicht berücksichtigt wurden, die aber zur Verbesserung der Wirksamkeit des Programms hilfreich sein können.

Die Evaluation kann also angewendet werden auf:

- ✓ Das Identifizieren von Bedürfnissen der Schüler zu Beginn des Programms.
- ✓ Sie ermöglicht die Beobachtung der Schülerleistungen während der Durchführung des Programms.
- ✓ Ermöglicht die Bewertung der von den Schülern erreichten Ziele.

Die Evaluierung ermöglicht es auch, die getätigten Investitionen im Kontext der Wirksamkeit zu begründen. Dieser Aspekt ist essentiell, wenn man davon ausgeht das Projekte und Programme durch öffentliche Gelder finanziert werden.

Eine klare und transparente Evaluation, die die Effizienz und Wirksamkeit eines Programms belegt, erleichtert es also, die Unterstützung öffentlicher Stellen zu erreichen.

1.3 EVALUATIONSPLAN

Das Verfahren von Programmevaluationen hängt in hohem Maße von den Merkmalen des Programms selbst sowie vom Kontext ab. Es wirkt sich somit auf das gewählte Evaluationsmodell aus, das als Referenz herangezogen wird. Am wahrscheinlichsten ist es jedoch bei den meisten Bewertungsverfahren.

- ✓ Ansatz
- ✓ Design
- ✓ Informationssammlung
- ✓ Analyse
- ✓ Bericht

Um eine Bewertung durchführen zu können, müssen vorher einige Fragen geklärt werden

Stellen Sie sich vor, Sie werden um eine Evaluation eines in Ihrem Arbeitskontext stattfindenden Programms gebeten. Dabei können Fragen auftauchen, wie:

- ✓ Wer soll die Evaluation durchführen?
- ✓ Warum wollen Sie das Programm bewerten?
- ✓ Was genau soll evaluiert werden (das gesamte Programm oder einen Teil davon)?
- ✓ Welche Neben- und Randbedingungen gibt es?
- ✓ Kann die Evaluation unter den aktuellen Bedingungen durchgeführt werden?

Weitere Fragen, die zu Beginn auftreten können, sind:

- ✓ Welche Elemente müssen beachtet werden?

- ✓ Gibt es Indikatoren zum Handeln?
- ✓ Welche Qualitätskriterien sollten verwendet werden?
- ✓ Welche Informationen stehen über den Kontext des Programms hinaus zur Verfügung?

Dies sind einige der Punkte, die vor Beginn der Evaluierung geklärt werden müssen.

Realisierung des Programms

Bei der Realisierung eines Programms können wir zwei Situationen beschreiben. Zum einen, dass das Programm selbst die Evaluation geplant hat bzw. beinhaltet und zum anderen, dass dies nicht der Fall ist.

Im ersten Fall wird es möglich sein, eine Evaluation ausgehend von den im eigentlichen Programm festgelegten Grundsätze umzusetzen. Diese Evaluation kann durch ein allgemeineres Konzept ergänzt oder verbessert werden, das zwar nicht im Programm vorgesehen ist, jedoch zur Verbesserung der Bewertung angewendet werden kann. Im zweiten Fall sollte die Evaluation auf den Merkmalen und der Entwicklung des Programms beruhen, Obwohl sie nicht eigentlicher Bestandteil des Programmes selber ist.

Angenommen, ein Programm wird in der direkten Nachbarschaft gestartet, um die Müll-recyclingrate zu erhöhen. Im selben Programm gab es eine Phase über die Folgenabschätzung und -verbesserung, die mit dem Programm erzielt werden kann. In dieser Phase wird

... nach der Anwendung des Programms eine Erhöhung der Recyclingquote um 10% erwartet. Das könnte als wenig erfolgreich angesehen und in anderen Stadtteilen sogar als Fehlschlag gewertet werden, da die betrachtete Nachbarschaft eine eigene Charakteristik hat. Allerdings können in diesem Fall 10% als Erfolg gewertet werden, da im betrachteten Stadtteil, der sich im Spannungsfeld langwierig bestehender sozialen Spannungen bewegt, eine solche Verbesserungen urbaner Stadtentwicklung noch nie erreicht wurden.

Mit einer solchen Interpretation setzt der Evaluator den Prozess die Wirksamkeitskriterien in einen entsprechenden Kontext, der auch von Aspekten außerhalb des Programms beeinflusst wird. Wenn dieser Kontext ausbleibt, können vollkommen verzerrte und unrealistische Bewertungen entstehen. Deshalb ist das Programm im Vorteil, in dessen eigentlichen Setup die Evaluation enthalten ist.

Unabhängig davon, ob es im Programm enthalten ist oder nicht, muss bei der Durchführung einer Evaluierung immer berücksichtigt werden, welche Lösungen vorgeschlagen und welche Alternativen dazu genannt werden. Damit besteht die Möglichkeit, den jeweils effektiveren Weg auszuwählen.

Der Erfolg einer Evaluation

Der Erfolg einer Evaluierung hängt in hohem Maße von der Prozessplanung und -organisation ab. Essentiell ist, dass in der Planung das Sammeln von Informationen berücksichtigt wird. Das Evaluierungsteam muss so bald wie möglich über die Datenerfassungstermine (sowohl Initialen als auch gegebenenfalls später) verfügen und sicherstellen, dass das Verfahren zur Erfassung der Daten (Fragebögen, Beobachtungen, Tests usw.) angewendet wird und entsprechende Autorisierungen dafür vorliegen.

Nach dem Sammeln der Informationen verfügt der Bewerter über sensible Daten, die der Verschwiegenheit unterliegen. Wie bei jeder anderen Art von personenbezogenen Daten unterliegen auch die Daten, gesetzlichen Datenschutzbestimmungen sowie gewissen Ethik-Codecs von Evaluatoren. Ein adäquates Speichersystem zusammen mit der absoluten Vertraulichkeitsgarantie sind zwei wesentliche Anforderungen, die der Evaluator in Bezug auf die personenbezogenen Daten benötigt.

Unmittelbar nach Erhalt der Informationen wird mit der Analyse begonnen. Dies geschieht unter Einhaltung der Bewertungsziele und der Anforderungen an die Evaluation. Die Art der Analyse und die verwendeten technischen Ressourcen hängen von den Merkmalen der Daten, den Merkmalen des Programms, den für die Bewertung verfügbaren Ressourcen usw. ab.

In einer Programmevaluation können die erhobenen Daten sowohl qualitativ (textuell) als auch quantitativ (numerisch) sein. Ebenso ist es möglich, die Analyse qualitativ (Bewertung der Informationsqualität) oder quantitativer (Bewertung der ermittelten Maßnahmen) durchzuführen. Wenn beide Möglichkeiten gekreuzt sind, werden vier Analyse-kategorien benannt:

- ✓ qualitative Analyse qualitativer Daten
- ✓ qualitative Analyse quantitativer Daten
- ✓ quantitative Analyse qualitativer Daten
- ✓ quantitative Analyse quantitativer Daten

Reporting

Schließlich muss das Bewerterteam alle seine Arbeiten in den entsprechenden Bericht übersetzen. Abhängig von den wissenschaftlichen Anforderungen an die Evaluatoren oder anderer Zielgruppen, für die der Bericht erstellt wird, wird dieser in dem einen oder anderen Format verfasst. Es wird demnach nicht ein und derselbe Bericht für bspw. die Eltern von Schülern, Lehrergremium, Supervisionsteam oder etwa Politiker des Bildungsministeriums verfasst. Trotz der unterschiedlichen Möglichkeiten, den Bericht zu übernehmen und auch der unterschiedlichen Ausdrucksstile, lassen sich die Berichte in einer allgemeinen Struktur zusammenfassen, die ein Deckblatt, eine Zusammenfassung, Informationen zum evaluierten Programm, die Evaluierungsmethode und die erzielten Ergebnisse enthält. Die Diskussion, die sich aus den Ergebnissen ableiten lässt, endet mit einer Reihe von Schlussfolgerungen und Empfehlungen.

★ Übung!

Erzählen Sie es Ihren Eltern (oder Ihrem Partner). Stellen Sie sich vor, wie Sie gefragt werden, was Sie in diesem Kurs tun. Dabei haben Sie sich doch bestimmt die Idee, ihnen von einer Programmevaluierung zu erzählen, oder? Wenn Sie das tun, wird er / sie Sie anstarren und fragen, bitte erkläre genauer, was eine Evaluation ist.

Versuchen Sie so zu erklären, als hätte er / sie noch nie von Evaluationen gehört und auch sonst keine Beziehung dazu. Notieren Sie die Erklärung.

2 BASISKONZEPTE

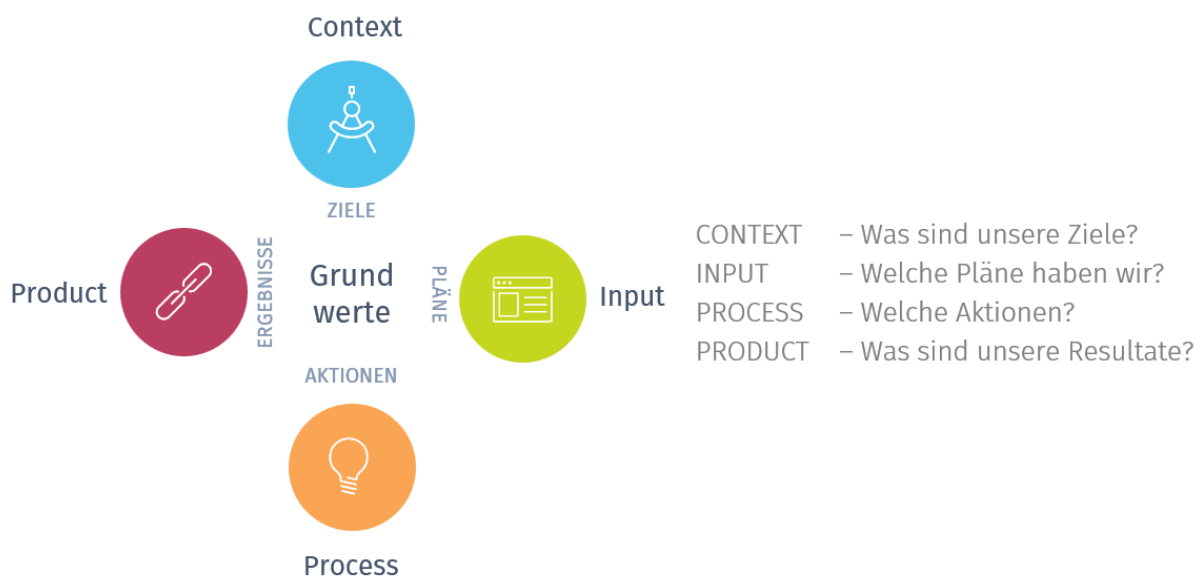
Viele Konzepte und Strategien aus unterschiedlichen Quellen werden in akademischen Bereichen der Programmevaluierung verwendet. In diesem Abschnitt werden wir einige von ihnen aufzeigen und durch Aufgaben und Übungen greifbar machen.

2.1 CIPP MODEL

Das CIPP-Modell von Stufflebeam ist ein äußerst nützlicher Ansatz, um die Evaluierung von Programmen zu organisieren.

Alles, was wir im CIPP-Modell über **CONTEXT**, **INPUT**, **PROCESS**, **PRODUCT**, Aktionen, Ziele und Ergebnisse zu sagen haben, basiert auf den Grundwerten unserer Bildungsorganisation und bezieht sich darauf.

Es ist die Aufgabe des Managements und des Vorstands einer Bildungsorganisation, gemeinsam mit dem Team und den Stakeholdern die Grundwerte zu formulieren. Für INTENSE finden sich die Grundwerte in der **CHARTA**, dem **TRAININGSPROGRAMM** für Transition Coaches, **DISKUSSIONEN** mit Stakeholdern und dem **TRANSITION PLAN** für die Organisation. Siehe unten.



Um mehr Einblick in die Anwendung des Modells zu bekommen empfehlen wir den folgenden Artikel aufmerksam durcharbeiten:

<http://openjournals.libs.uga.edu/index.php/jheoe/article/download/628/482>

★ Übung!

Suchen Sie weitere Informationen zum CIPP-Modell im Internet. Erstellen Sie eine eigene kurze Zusammenfassung des CIPP-Modells.

2.2 CONTRIBUTION ANALYSIS

“Questions of cause and effect are critical to assess the performance. When it is not practical to develop a project to assess the performance, contribution analysis can provide credible assessments of cause and effect. Verifying the theory of change that the programme is based on and paying attention to other factors that may influence the outcomes, provides reasonable evidence about the contribution being made by the programme.”

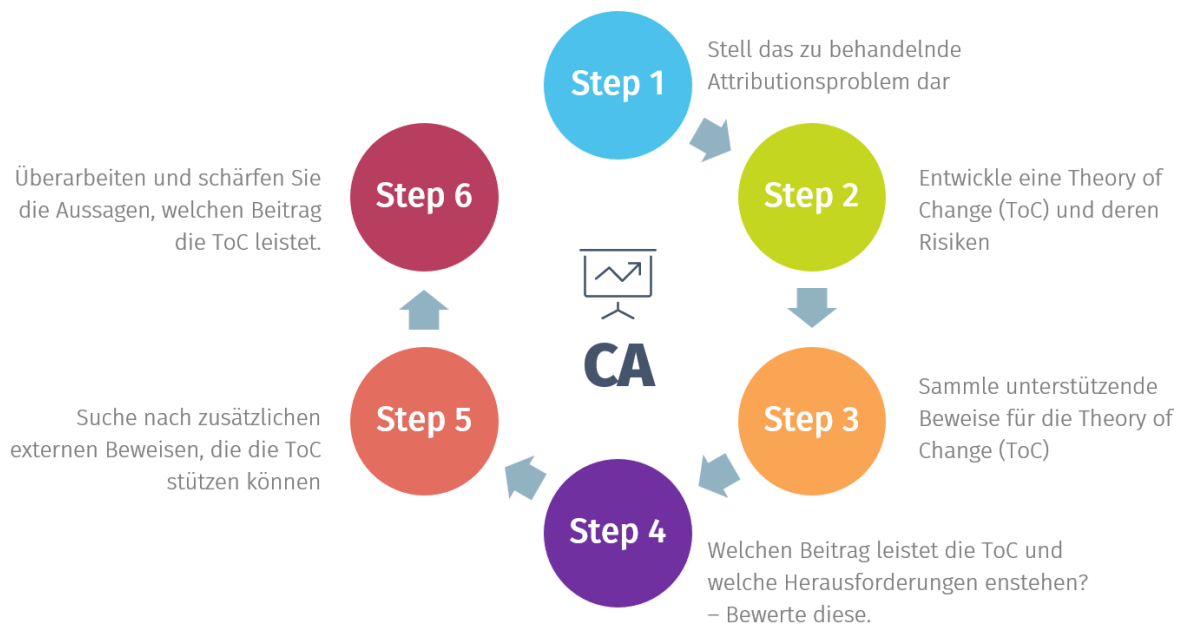
John Mayne

Contribution Analysis: An approach to exploring cause and effect. 2008

Es ist nicht beabsichtigt, eine direkte Kausalität zu beweisen. In der komplexen sozialen Welt, in der wir arbeiten, ist es unwahrscheinlich, dass eine Intervention den einzigen Einfluss auf bestimmte Ergebnisse hat. Daher müssen wir ein **ATTRIBUTION PROBLEM** formulieren. Inwieweit werden Ergebnisse aufgrund von Programmaktivitäten beobachtet?

Die Contribution Analysis ist ein strukturierter Ansatz, um zu verstehen, inwieweit beobachtete Ergebnisse eine Konsequenz einer bestimmten Intervention sind, im Vergleich zu anderen Faktoren. In der komplexen sozialen Welt, in der wir arbeiten, ist es unwahrscheinlich, dass eine Intervention den einzigen Einfluss auf bestimmte Ergebnisse hat.

Die Beitragsanalyse konzentriert sich auf das Verständnis, warum die beobachteten Ergebnisse aufgetreten sind, welche Rolle die Aktivität spielt und auf andere interne und externe Faktoren. Indem Sie Beweise für eine Argumentationslinie (eine Theory of Change) sammeln und vorlegen, können Sie eine vernünftige Schlussfolgerung ziehen. Es gibt sechs sich wiederholende Schritte in der Contribution Analysis. Jeder Schritt baut die **CONTRIBUTION STORY** auf und legt die Schwachstellen in der vorherigen Phase offen. Siehe unten.



★ **Step 1. Stellen Sie das zu behandelnde Attributionsproblem dar.**

Dieser Schritt ist notwendig, um die verschiedenen Fragen zu formulieren.

- ✓ Bestätigen Sie das Problem. Ein fundamentaler Fehler ist, dass wir andere oft scharf beurteilen, während wir uns selber nicht an den gleichen Maßstäben messen, indem wir unser eigenes (unethisches) Verhalten rationalisieren. In dieser Analyse muss man sich immer diesen Fehlers vor Augen führen.
- ✓ Bestimmen Sie die spezifische Ursache-Wirkung-Frage, die damit adressiert wird.
- ✓ Bestimmen Sie das erforderliche Vertrauensniveau.
- ✓ Informieren Sie sich über die Art des erwarteten Beitrags
- ✓ Bestimmen Sie die anderen wesentlichen Einflussfaktoren
- ✓ Beurteilen Sie die Plausibilität des erwarteten Beitrags in Bezug auf den Umfang des Programms

★ **Step 2. Entwickeln Sie die Theory of Change. (ToC)**

Dieser Schritt zusammen mit der **RESULTCHAIN** sind die Schlüsselinstrumente der Analyse. Die ToC und die Ergebniskette stehen also in direktem Zusammenhang zueinander.

Jede Intervention mittels eines Projekts (INTENSE) basiert auf Annahmen darüber, was das Projekt (durch Interventionen) bewirken wird. Eine Theory of Change ist also eine Erklärung dafür, warum wir glauben, dass Ihr Projekt zu positiven Ergebnissen führen wird.

Wenn eine Organisation eine Theory of Change erstellen möchte, muss man sich einige Gedanken machen:

1. Was sind die Voraussetzungen, um die Ziele von INTENSE zu verwirklichen?
2. Was sind die Risiken?

Darstellung der ToC:

- ✓ Bestimmen Sie den Detaillierungsgrad
 - ✓ Bestimmen Sie den erwarteten Beitrag des Programms
 - ✓ Führen Sie die Annahmen auf, die der ToC zugrunde liegen
 - ✓ Berücksichtigen Sie auch andere Faktoren, die das Ergebnis beeinflussen können
 - ✓ Bestimmen Sie, wie viel die ToC angefochten wird
3. Wie kann eine Kausalkette in Bezug auf die ToC realisiert werden?
 4. Formulieren Sie die Kausalkette in Bezug auf Inputs - Aktivitäten - Outputs - (Zwischen-) Ergebnisse und Endeffekte.

★ **Step 3: Sammeln Sie unterstützende Beweise für die Theory of Change.**

Bewerten Sie die Logik der Verknüpfungen in der ToC. Die Prüfung der Stärken und Schwächen der Logik, der Plausibilität der verschiedenen Annahmen in der Theorie und des Ausmaßes, in dem sie bestritten werden, liefern Hinweise, wo konkrete Beweise am dringendsten benötigt werden. Sammeln Sie die Beweise. Die Bestätigung der ToC ist zwingend in drei Bereichen erforderlich: Beobachtete Ergebnisse, Annahmen zur ToC und andere Einflussfaktoren.

- ✓ Belege für Ergebnisse und Aktivitäten
- ✓ Belege für Annahmen
- ✓ Hinweise auf andere Einflussfaktoren

★ **Step 4: Welchen Beitrag leistet die ToC und welche Herausforderungen entstehen? – Bewerten Sie diese.**

Die bisher entwickelten Beiträge zur Zielerreichung können nun zusammengestellt und kritisch bewertet werden. In dieser Phase sind noch weitere Fragen zu stellen. Einige wichtige Beispiele:

- ✓ Wie glaubwürdig ist das Konstrukt insgesamt?
- ✓ Welche Glieder in der Ergebniskette sind stark?
- ✓ Stimmen die Stakeholder dem Konzept / Konstrukt zu?
- ✓ Ist klar, welche Ergebnisse erzielt wurden?
- ✓ Sind wichtige Annahmen validiert?

★ **Step 5. Suchen Sie nach zusätzlichen, externen Beweisen.**

- ✓ Identifizieren Sie, welche neuen Daten benötigt werden.
- ✓ Justieren Sie die ToC.
- ✓ Sammeln Sie weitere Beweise.

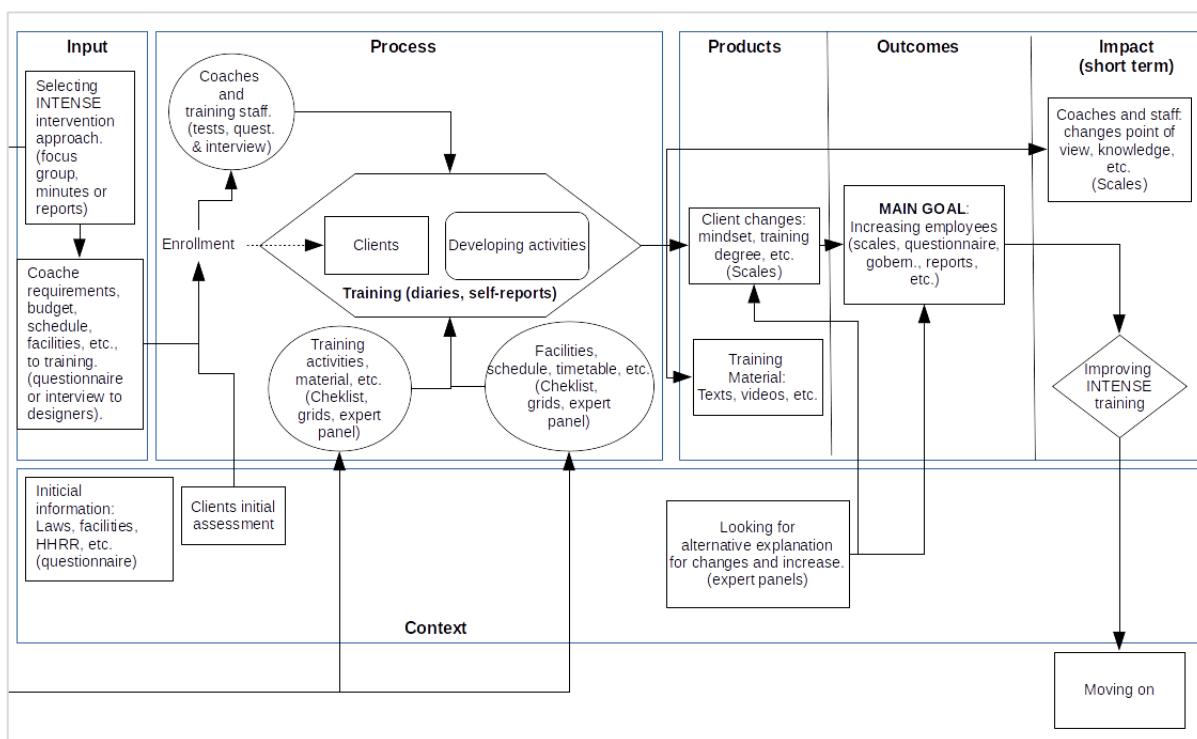
★ **Step 6. Hinterfragen Sie und schärfen Sie den Blick auf die Beiträge zur Zielerreichung (ToC).**

- ✓ Neue Beweise werden ein glaubwürdigeres Konstrukt schaffen.

- ✓ Unterstützung der schwächeren Teile der früheren Version oder Änderungsvorschläge für die aktuelle ToC.
- ✓ Es ist unwahrscheinlich, dass die überarbeitete ToC narrensicher sein wird, aber sie wird stärker und glaubwürdiger.

Die Contribution Analysis funktioniert am besten als sich wiederholender Kreislauf. Dabei ist es möglich, zu einen bestimmten Punkt im Kreislauf zurückzukehren und ggf. die Stärken und Schwächen der Beiträge neu bewerten.

Einen differenzierten Einblick in die ToC des Projektes INTENSE liefert die nachfolgende Grafik. Sie ist zentraler Bestandteil des Evaluationsdesigns und damit Basis ToC für alle individuell entwickelten ToCs der Partner.



Quelle: University of Malaga

Um Informationen und Daten zu sammeln, können abhängig vom Partner, der sie auswählt, die folgenden Instrumente verwendet werden:

- ✓ Offizielle Statistiken
- ✓ Offizielle Berichte
- ✓ Interviews
- ✓ Fragebögen oder Umfragen
- ✓ Skalen und Diagramme
- ✓ Auswertungen von Testverfahren
- ✓ Beobachtungen: Selbstbericht, Diaries, Fallstudien, Checklisten.

Für weiterführenden Input lesen Sie bitte die folgenden Artikel:

- ✓ <https://www2.gov.scot/Resource/Doc/175356/0116687.pdf>
- ✓ https://www.betterevaluation.org/sites/default/files/ILAC_Brief16_Contribution_Analysis.pdf

2.3 OECD KRITERIEN

Ein weiteres wichtiges Element sind die Kriterien, die als Leitfaden für die Durchführung der Evaluation dienen. Die Abteilung DAC-OECD hat eine Reihe von Kriterien erarbeitet, die als Aktivitäten vorgeschlagen werden. Bitte lesen Sie die Bewertungskriterien des DAC-OECD.

<https://www.oecd.org/development/evaluation/qualitystandards.pdf>

2.4 THEORY OF CHANGE

*“Every intervention in an organization is based on assumptions that together form a theory. The Theory of Change approach tests the theory of performers – because nothing is as practical as a good theory.”
(Erik Snel, 2013)*

Die unten gezeigte Ergebniskette visualisiert den seriellen Zusammenhang hinter der Theorie of Change (ToC), die lt. Wikipedia

“... explains the process of change by outlining causal linkages in an initiative, i.e., its shorter-term, intermediate, and longer-term outcomes. The identified changes are mapped – as the “outcomes pathway” – showing each outcome in logical relationship to all the others, as well as chronological flow. The links between outcomes are explained by “rationales” or statements of why one outcome is thought to be a prerequisite for another”.

Die ToC beschreibt parallel auch den Verlauf des Implementierungsprozesses. Der Plan den man daraus erstellen kann, ist sozusagen der Leitfaden für eine erfolgreiche Umsetzung. Er hilft auch, die Denkweise der Menschen in einer Organisation zu ändern und einen gemeinsam getragenen Prozess zu etablieren. Bekanntlich ändert sich ohne ein Umdenken, das in den Menschen stattfindet, nichts.

Ein erfolgreicher Umsetzungsprozess hängt von vielen Faktoren ab, wie angemessener Schulung, theoretischem Wissen, Unterstützung durch das Management, vor allem aber von einem klaren Bekenntnis zum Wandel. Wenn diese Verpflichtung nicht besteht, ist es fast unmöglich, INTENSE als Gesamtkonzept oder in Teilen umzusetzen.

RESULTS CHAIN



Für weiterführende Informationen und Input sehen Sie sich das folgende Video an:

https://youtu.be/KRptX_DNL2Q

★ Übung!

Antworten Sie kurz auf die Frage - Ihrer Meinung nach: Welche Rolle kann die Theory of Change in Ihrer Arbeit als Coach spielen?

3 EVALUATIONSDESIGN

3.1 INTRO

Ein Design ist ein Aktionsplan für die in der Evaluation auszuführenden Schritte, um Änderungen in dem bewerteten System zu identifizieren. In der Evaluierung können drei Szenarien angegeben werden, die bestimmen, welche Art von Design verwendet werden soll. Die ersten beiden basieren auf der Annahme, Gruppen miteinander zu vergleichen. Die dritte ist eine Alternative, wenn Vergleichsgruppen unmöglich sind:

ERSTES SZENARIO: Wiederholend-kontrollierte Studien. Dies sind Designs, die ausgeführt werden, wenn Sie die volle Kontrolle über den Start des Programms sowie über die teilnehmenden Personen haben. Dies beinhaltet die Steuerung über die Programmabschlusszeiten, die Auswahl der Teilnehmer, die Zuordnung zu verschiedenen Versionen oder Modalitäten des Programms, die von ihnen verwendeten Ressourcen usw.. In diesem Fall ist das Design typisch für den Bereich der experimentellen Forschung.

ZWEITES SZENARIO: Nicht zufällige Gruppendesigns. Wenn es nicht möglich ist, eine vollständige Steuerung über das Programm oder die Teilnehmer zu haben, muss die Kontrolle über frühere Designs aufgegeben werden. Es geht um Situationen, in denen Sie Evaluierungspläne entwickeln können, die in Form und Organisation den vorherigen ähnlich sind. Da jedoch keine vollständige Kontrolle besteht, sollten die Ergebnisse mit Vorsicht betrachtet werden. Mit anderen Worten, obwohl das Design dem ersten Szenario ähnelt, bedroht die Unmöglichkeit, alle Aspekte des Prozesses zu kontrollieren, die Gültigkeit der Ergebnisse. Aus diesem Grund sollten die genannten Ergebnisse eher kritisch bewertet und ergänzende Wege gefunden werden, um gesicherte Resultate zu erzielen.

DRITTES SZENARIO: Kohärenzstudien. Dabei handelt es sich um ein Szenario, in dem nicht garantiert ist, dass Kontrolle über die verschiedenen und bedeutenden Elemente besteht. Mit diesen Studien wird gezielt nach Alternativen gesucht, die unabhängig vom Vergleich der Maßnahmen zwischen 2 oder mehr Gruppen sind.

Diese drei Szenarien lassen sich anhand zweier grundlegender Ansätze organisieren:

KONTRAFAKTISCHE DESIGNS: Sie bestehen aus Entwürfen, die faktische Gruppen verwenden. Eine faktische oder kontrafaktische Gruppe ist die Gruppe, bei der das Programm nicht angewendet wird, die jedoch in Charakteristika genau der Gruppe entspricht, bei der es angewendet wird. Innerhalb dieser Designs sind wiederholend-kontrollierte Studien am wirkungsvollsten. In der differenzierten Anwendung können auch Teile als nicht zufällige Gruppendesigns organisiert werden, obwohl diese weniger Aussagekraft besitzen.

NICHT KONTRAFAKTISCHE DESIGNS: Das sind solche, bei denen man sich nicht auf gruppenbezogene Fakten verlassen kann und somit der Vergleich zwischen Gruppen nicht möglich ist. Zu dieser Art von Design gehören Kohärenzdesigns.

3.2 KAUSALE ATTRIBUTION

Eines der Grundelemente der Designs besteht darin, die kausale Attribution für das Programm zu erhalten. Um mehr zu erfahren, lesen Sie bitte das folgende Dokument:

<https://www.unicef-irc.org/publications/751-overview-strategies-for-causal-attribution-methodological-briefs-impact-evaluation.html>

★ Übungen!

1. Ein Programm zur Kontrolle der Aggressivität wird in einer Haftanstalt mit akzeptablen Ergebnissen durchgeführt. Trotzdem wurde beschlossen, ein weiteres auf gegenseitiger Achtsamkeit basierendes Programm zu starten. Um die Wirksamkeit dieses neuen Programms im Vergleich zu dem anderen zu überprüfen, wurde zufällig eine Gruppe von Teilnehmern ausgewählt. Geben Sie an, welches Evaluationsdesign angewendet werden kann und begründen Sie kurz Ihre Antwort.
2. Ein Englischlehrer verwendet eine neuartige Methode, die auf der Zusammenarbeit mit Lehrern anderer Fächer beruht. Zum Beispiel entwickeln sie mit dem Wirtschafts- und Sozialkundelehrer das Budget eines fiktiven Unternehmens auf Englisch. Stellen Sie sich vor, Sie müssen sich bei der Studentengruppe anmelden, um relevante Daten zu erheben. Der Kurs läuft ein ganzes Semester und Sie benötigen die Einwilligungen der Schulleitung, der Lehrer und der teilnehmenden Schüler. Machen Sie anhand dieser Angaben einen Designvorschlag für die Evaluierung, in dem berücksichtigt wird, dass mehr als ein Schüler für die Erhebung der Daten angesprochen werden kann. Begründen Sie Ihren Vorschlag kurz.
3. Sie arbeiten in der Einwanderungszentrale und versuchen, sich an die kürzlich genehmigten Regelungen anzupassen. Ihrer Wahrnehmung nach hat die Einführung dieser Regelung eine Verschlechterung der Aufmerksamkeit der Nutzer zur Folge. Daneben scheint die Bearbeitungszeit bis zur Validierung der Benutzerdokumentation erheblich zugenommen zu haben. Ebenso hat sich die Anzahl der Nutzer pro Sozialarbeiter um durchschnittlich 3 Personen erhöht. In Anbetracht dieser Situation geben Sie an, welches Evaluierungsdesign Sie anwenden könnten, um diese Situation faktisch zu untersetzen und begründen Sie Ihren Vorschlag.

4 INDIKATOREN

4.1 DEFINITION

Obwohl es verschiedene Ansätze für die Erklärung und Anwendung von Indikatoren gibt (Rodríguez Jaume, 2000), bezieht sich dieser aus pragmatischer Sicht auf eine Maßeinheit, die die Überwachung und Bewertung eines Systems ermöglicht. Wenn es um die Definition geht, scheint der Begriff mit anderen Konzepten verknüpft zu sein. Darunter das Messkonzept, die Variable oder der Index. Wir möchten hier einen transparenten Überblick zu diesen Begriffen bieten, damit ihre Abhängigkeiten und Überschneidungen erkennbar werden und wir künftig Verwechslungen ausschließen können.

- ✓ Variable: Allgemeiner Name, der sich auf die Wertemenge bezieht, die von einer Kennzahl übernommen werden kann. Beispiel: Einwohnerzahl.
- ✓ Messung: Numerische Zuordnung zu einer Realität oder einem Phänomen. Dies sind die typischen Daten, die als Zahl oder als Attribut dargestellt werden können. Beispiel: Anzahl der Einwohner in einem bestimmten Moment, die numerisch (z.B. 42.670) oder als Attribut (viele) ausgedrückt werden kann.
- ✓ Indikator: Maßnahme, die eine bestimmte Situation in einem Programm abbildet. Seine Funktion ist es, die Entscheidungsfindung zu unterstützen. Beispiel: Die Einwohnerzahl zu einem bestimmten Zeitpunkt beträgt x (z.B. 1.243).
- ✓ Index (oder Referenz): Gruppe von Indikatoren, die in einem Ausdruck zusammengefasst sind. Er baut mit der Kombination mehrerer Indikatoren auf. Beispiel: Wachstumsrate ($CI = x(t-1) / x$)
- ✓ Standard (oder Rubrik): Reihe von Indizes oder Indikatoren, die Informationen zur Leistung liefern. Beispiel: Der IC ist zu einem bestimmten Zeitpunkt sehr hoch, hoch, mittel oder niedrig.

Wie ersichtlich ist, ist jeder Begriff eng mit dem Folgenden verbunden. Diese Verbindungen werden in der Tabelle deutlich.

Variable ↓	Phänomen oder Dimension, die unterschiedliche Werte annehmen kann
Messung ↓	Zuordnung eines (numerischen) Wertes zum interessierenden Phänomen
Indikator ↓	Maß oder konkreter, spezifischer Wert, bei dem eine Variable zu einem bestimmten Zeitpunkt gemessen wird
Index ↓	Aggregation von Indikatoren
Standard	Bewertung (subjektiv) des Phänomens basierend auf den Ergebnissen des Indikators (oder der Indikatoren)

Table: Beziehung zwischen den Begriffen

4.2 CHARAKTERISTIKA VON INDIKATOREN

Die Indikatoren müssen nützlich und gültig sein. Das heißt, sie müssen dem in der Untersuchung festgelegten Zweck dienen und sich auf die Informationen beziehen, die gesammelt werden sollen.

Wenn Sie beispielsweise die Anzahl der Benutzer des öffentlichen Gesundheitssystems wissen möchten, muss der Indikator die Anzahl der Benutzer sein, die das Gesundheitssystem tatsächlich nutzen und nicht etwa die Anzahl der mit dem System verbundenen Personen im Gesundheitswesen.

Zusätzlich zu diesem Grundmerkmal wird in der Literatur häufig auf die sogenannten SMART-Kriterien verwiesen. Die SMART-Kriterien wurden ursprünglich für das Verfassen von Zielen (Doran, 1981) sowohl für die Forschung als auch für Organisationen, Projekte oder Programme vorgeschlagen. Derzeit sind sie jedoch auf andere Aspekte des wissenschaftlichen und organisatorischen Gebiets verallgemeinert worden. SMART ist die Abkürzung für spezifisch, messbar, erreichbar, ergebnisorientiert und zeitlich begrenzt.

- ✓ Spezifisch: Ein Indikator ist spezifisch, wenn er klar und präzise ausgedrückt wird
- ✓ Messbar: Der Indikator sollte so ausgedrückt werden, dass leicht zu erkennen ist, wie die Größe bestimmt wird.
- ✓ Erreichbar oder realisierbar: In der Regel wird als realisierbar verstanden, dass der Indikator für das zu evaluierende Programm geeignet ist. Man kann es auch so verstehen, dass der Indikator so ausgelegt wird, dass seine Verwendung möglich ist.
- ✓ Ergebnisorientiert: Der Indikator muss relevant sein, um messen zu können. Die Datenerhebung muss mit den Informationen übereinstimmen, die analysiert werden sollen
- ✓ Zeitlich begrenzt: Der Indikator muss innerhalb eines bekannten Zeitraums wirken.

Trotz der Nützlichkeit dieser Kriterien ist es klar, dass sie aus dem Festlegen von Zielen stammen. Eine Überprüfung ist in jedem Fall erforderlich, um eine bessere Anpassung an den Bewertungsbereich zu erreichen.

4.3 TYPEN VON INDIKATOREN

Die Indikatoren können nach unzähligen Formen klassifiziert werden: nach der Komponente der Bewertung, dem Grad der Objektivität, der Einfachheit, der Relativität, intern oder extern usw. Letztlich hängen die Indikatoren nur von der Komponente, in der sie verwendet werden, ab:

Kontextindikatoren: Das sind alle, die Informationen über die Situation sammeln, in der das Programm stattfindet, z.B. die Anzahl der Personen, die vom Programm profitieren, das wirtschaftliche Niveau der interessierenden Bevölkerung, die Anzahl der behandelten Patienten usw.

Input: Indikatoren, die unter anderem die am Programm beteiligten Humanressourcen, die wirtschaftlichen, materiellen und soziokulturellen Aspekte messen. Dazu gehören auch die Indikatoren, die Daten über die Programmgestaltung und den Ausführungsplan sammeln (wirtschaftliches Budget, geplanter Kalender, Anzahl der geplanten Sitzungen usw.).

Prozess: Die hier verwendeten Indikatoren versuchen Informationen darüber zu sammeln, wie sich das Projekt in der Realität entwickelt. Beispiel: Anzahl der Schulungen, Bewertung des Materials durch die Benutzer, Stunden für eine Aufgabe, Personen, die in einer bestimmten Aktivität tätig sind, Kosten für die Logistik des Programms usw.

Produkte: Diese Kriterien versuchen, Daten über die Auswirkungen des Programms zu sammeln. In dieser Komponente müssen wir die Produkte von Ergebnissen und Auswirkungen unterscheiden. Bei den Produkten handelt es sich um die unmittelbarsten Auswirkungen oder kurzfristigen Ergebnisse, z.B. Anzahl der gebildeten Personen, endgültige Kosten des Programms, durchschnittliche Benutzerbewertung am Ende eines Programms usw. Dies sind die Zwischenergebnisse. Dies wirkt sich beispielsweise auf die Änderung von Vorschriften an einem Ort aus, die ganz oder teilweise auf die Produkte des Programms zurückzuführen sind.

Letztendlich ist der Impact die langfristige Auswirkung eines Programms, einschließlich der geplanten und ungeplanten positiven und negativen Auswirkungen. Zum Beispiel die Verringerung der Umweltverschmutzung in einer Region, die Änderung der Einstellung eines Teils der Bevölkerung in Bezug auf ein Problem, den Prozentsatz des Verbrauchs zuckerfreier Produkte.

★ Übung!

Die Europäische Union verfügt über den Eurostat-Webpace, auf dem eine große Menge an Daten öffentlichen bereitgestellt werden. Erstellen Sie einen kurzen Aufsatz mit maximal 1000 Wörtern über die projizierte Bevölkerungsentwicklung in der Europäischen Union. Dazu müssen Sie unter den in Eurostat gesammelten Indikatoren nach dem entsprechenden Indikator dafür suchen.

Link: <https://ec.europa.eu/eurostat/data/database>

5 BERICHTE, KOMMUNIKATION UND ETHISCHE ANFORDERUNGEN

5.1 BERICHTE

Die Berichte sind ein Ergebnis der Evaluation. Die Evaluation muss folgendes sicherstellen:

- ✓ Antragsteller, interessierte Vertreter und andere Empfänger der Berichte sollten genügend Informationen über die Entwicklung und die Auswirkungen des Programms erhalten.
- ✓ Antragsteller und interessierte Beauftragte nutzen die Berichte, um die Programmentwicklung, ihren Wert und die damit verbundenen Entscheidungen umfassend zu verstehen.

Beim Ausarbeiten der Berichte sollte berücksichtigt werden, dass die Auftraggeber und Administratoren des Programms meist keine lange Berichte mögen. Sie suchen vor allem nach Berichten, bei denen es leicht ist, die für Entscheidungen erforderlichen Informationen rasch zu finden.

Das bedeutet nicht, dass Informationen fehlen. Berichte sollten auch die technischen Aspekte des Bewertungsprozesses enthalten. In jedem Fall hängen die Art des Inhalts und die Art der Verbreitung von der Art des Empfängers ab. Kurz gesagt, eine ausgewogene Balance zwischen Inhalt, Umfang und Interesse des Empfängers ist das oberste Ziel.

Dabei ist zu beachten, dass Berichte auch ein Mittel zur Erzielung von Transparenz im Bewertungsprozess sind, was natürlich dessen Zuverlässigkeit und Nützlichkeit unterstreicht.

Es gibt zwar viele allgemeine Regeln, wie Berichte zu verfassen sind – wir möchten an dieser Stelle einen Vorschlag für einen vollständigen Evaluierungsbericht unterbreiten.

Allgemeine Charakteristiken:

- ✓ Schreiben Sie maximal 30 bis 50 Seiten (mit separaten Anhängen). Die technischen Anhänge haben wenig operativen Nutzen und können in kurzen Berichten auch weggelassen werden.
- ✓ Fügen Sie am Anfang eine kurze Zusammenfassung ein (maximal 1 Seite).
- ✓ Fügen Sie auch eine Zusammenfassung mit den wichtigsten Ergebnissen und Empfehlungen bei (maximal 4 Seiten).
- ✓ Programmbeschreibung.
- ✓ Beschreibung des Evaluationsverfahrens.
- ✓ Einzelheiten zu den Analysen und Instrumenten können den Anhängen beigelegt werden.

Zu den Abschnitten:

- ✓ Einführende Zusammenfassung
- ✓ Beschreibung des Programms: Ziele, Ressourcen, Realisierungsprozess, Ergebnisse und Schlussfolgerungen einiger früherer Bewertungen.
- ✓ Beschreibung der Evaluation: Ziele und Schlüsselfragen; Design; Indikatoren, Variablen und Instrumente; verarbeiten; und Ergebnisse.
- ✓ Schlussfolgerungen und Empfehlungen.
- ✓ Referenzen.
- ✓ Anhänge.

5.2 KOMMUNIKATIONSSTRATEGIE

Im Allgemeinen werden häufig mehrere Strategien und Mittel eingesetzt, um die Ergebnisse der Bewertung zu verbreiten, was bedeutet, dass mehrere Berichte verfasst werden. In dieser Hinsicht hängt die Art der Kommunikationsmittel weitgehend vom Empfänger ab. Ganz allgemein müssen Sie mindestens drei mögliche Zielgruppen berücksichtigen: (1) Techniker und Administratoren, (2) politische Entscheidungsträger und Manager und (3) Fachleute.

Für jede dieser Personenkreise wird ein spezifischer Kommunikationsplan erstellt, der die zu verwendenden Mittel (Berichte, Konferenzen, Artikel in Fachzeitschriften, Arbeits-tische usw.) sowie einen Verteilungskalender enthält.

Berichte werden in der Regel offen für interessierte Zielgruppen präsentiert. Daher wird mindestens eine Veranstaltung für die Übermittlung, Präsentation oder Veröffentlichung jedes Berichts durchgeführt. Dies verleiht dem Prozess zusätzlich zu dem Werbeeffect Transparenz.

Schließlich ist ein sehr interessantes Thema mit den für die Verbreitung zu verwenden- den Ressourcen verbunden. Die Rolle der gegenwärtigen Globalisierung, das Potenzial neuer Technologien sowie die Auswirkungen sozialer Netzwerke und die Anzahl möglicher Ressourcen haben in den letzten Jahren exponentiell zugenommen und müssen berücksichtigt werden. Einige der möglichen Ressourcen sind die folgenden:

- ✓ Kurzvorträge.
- ✓ Informationshinweise, in denen die Schlussfolgerungen und Empfehlungen zusammengefasst sind. Sie können das Newsletter-Format verwenden, das sowohl physisch als auch digital veröffentlicht wurde.
- ✓ Blogs, in denen der Prozess, die Ergebnisse, Schlussfolgerungen usw. erläutert werden.
- ✓ Ausführliche Papierberichte mit allen Aspekten der Bewertung, die an das Auftraggeberinstitut geliefert wurden.
- ✓ Pressemitteilungen.
- ✓ Wissenschaftliche Artikel für Fachzeitschriften.
- ✓ Kurze Teilberichte mit verschiedenen spezifischen Aspekten der Bewertung.

5.3 ETHISCHE ANFORDERUNGEN

Zusammenfassend sind die wichtigsten ethischen Normen für die Personen, die eine Evaluation durchführen müssen, folgende:

- ✓ Verursachen Sie keine Schäden bei Personen! Dies ist das Hauptrisiko bei der Entwicklung der Evaluation. Schaden wird als körperlicher Zustand (Wunden, Schmerz, Unbehagen usw.) als psychischer Zustand (Angst, Unwohlsein, Stress, Angst usw.) als sozialer Zustand (schlechtes soziales Image, Konfrontation mit Gleichgestellten usw.) als wirtschaftlicher oder irgendeiner Art verstanden andere Art.
- ✓ Schließen Sie Personen nicht von Interventionen aus, von denen bekannt ist, dass sie von Nutzen sind. Dies bedeutet, dass die Bewertung keinen Benutzer daran hindern sollte, von einem Programm zu profitieren, wenn es sich als nützlich erwiesen hat.
- ✓ Die Zuordnung zu Vergleichsgruppen - nicht nach Bewertungskriterien - sollte zu Beginn des Programms umgesetzt werden. Es ist eine Anforderung an die Planung der Evaluierung. Die Bewerter könnten versucht sein, die Auswahlkriterien für die Teilnehmer zu ändern oder die Auswahl und Zuordnung nach nicht methodischen Kriterien vorzunehmen.
- ✓ Effizientes Management und Schutz privater und personenbezogener Informationen.

Eine andere wichtige ethische Dimension konzentriert sich auf die Fragen der Transparenz, d.h. auf alle Personen, die wissen möchten, wie eine Evaluation durchgeführt wurde. Hierzu muss die Evaluation folgenden Kriterien genügen: Ausschreibend, Unparteilichkeit und Reproduzierbarkeit.

★ *Quellen und weiterführende Links*

DAC (1991). Principles for Evaluation of Development Assistance. Paris: OECD. Disponible en <http://www.oecd.org/dac/evaluation/50584880.pdf>.

Doran, G. (1981). "There's a S.M.A.R.T." way to write management's goals and objectives, *Management Review*, 70(11), 35-36.

Peersman, G. (2014). Criterios de evaluación. Síntesis metodológicas: evaluación de impacto nº3. Florencia: Centro de Investigaciones UNICEF.

Pornpitakpan, C. (2004). The persuasiveness of source credibility: A critical review of five decades' evidence. *Journal of Applied Social Psychology* 34(2):249-281.

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: a systematic approach* (7ª Ed.). Londres: SAGE.

Stufflebeam, D. L., & Zhang, G. (2017). *The CIPP Evaluation Model: How to Evaluate for Improvement and Accountability*. Nueva York: Guilford Publications.

https://www.betterevaluation.org/en/evaluation-options/interim_reports

https://www.usaid.gov/sites/default/files/documents/1870/How-to-Note_Preparing-Evaluation-Reports.pdf

http://www.excellenceforchildandyoung.ca/sites/default/files/docs/webinars/_attach/evaluation_framework_template_example.pdf

https://www.volunteerflorida.org/wp-content/uploads/2017/02/example_process_evaluation_plan_annotated-2.12.15.pdf

https://wmich.edu/sites/default/files/attachments/u350/2017/eval-report-content-checklist_0.pdf

https://www.betterevaluation.org/en/blog/producing_engaging_accessible_evaluation_reports

https://www.betterevaluation.org/en/rainbow_framework/report_support_use/develop_reporting_media

https://www.betterevaluation.org/evaluation-options/interim_reports

https://www.betterevaluation.org/evaluation-options/final_reports

