
Leitfaden für die Planung und Durchführung von Wirkungsevaluationen



Wädenswil und Winterthur, 20. Januar 2025

Impressum

Herausgeber

Bundesamt für Umwelt (BAFU)

Das BAFU ist ein Amt des Eidg. Departements für Umwelt, Verkehr, Energie und Kommunikation (UVEK)

Autor*innen

Urs Müller, Silvia Burgdorf und Sarah Vaclavik

Forschungsgruppe Nachhaltigkeitskommunikation und Umweltbildung ZHAW

Sandra Wilhelm

anders kompetent GmbH

Begleitgruppe BAFU

Andrea Bader, Beat Bringold und Lilith Wernli

Zitierung

Müller, U., Wilhelm, S., Burgdorf, S., Vaclavik, S. (2025): Hilfestellung für selbständig durchgeführte Wirkungsevaluationen. Bern: Bundesamt für Umwelt BAFU.

Titelbild

Jacob Lund, Adobe Stock

PDF-Download unter

Eine gedruckte Fassung kann nicht bestellt werden.

Dieser Leitfaden wurde im Auftrag des Bundesamtes für Umwelt (BAFU) verfasst. Für den Inhalt ist allein der Auftragnehmer verantwortlich.

Inhalt

EINLEITUNG.....	4
WIRKUNGSMODELL – OHNE ZIEL KEIN WEG.....	5
1.1 DAS WIRKUNGSMODELL	5
1.2 DIE BEGRIFFLICHKEITEN.....	6
OUTCOMES – WIRKUNGEN FORMULIEREN	7
2.1 EIN «GUTER» OUTCOME...	7
2.2 GOOD PRACTICE-OUTCOMES.....	8
INDIKATOREN – WIRKUNGEN GREIFBAR MACHEN.....	9
3.1 INDIKATOREN ENTWICKELN	9
3.2 GOOD PRACTICE-INDIKATOREN	10
METHODEN – WIRKUNGEN MESSEN	11
4.1 VORÜBERLEGUNGEN ZUR METHODENWAHL.....	11
4.2 AUSWAHL AN METHODEN	12
4.3 GOOD PRACTICE-METHODEN.....	17
WEITERFÜHRENDE LITERATUR	19
QUELLEN	20

Einleitung

«Wir wollen etwas bei unseren Teilnehmenden bewirken», so bekunden wohl alle Anbietenden von Umweltbildungsformaten ihr Engagement. Und gerne möchten sie auch wissen, was genau sie durch ihre Projekte bei ihren Zielgruppen bewirken. Doch Wirkungsevaluationen stellen für viele Anbietende mehr Last als Reiz dar. Sie werden gemacht, wenn sie verlangt sind. Das sollte nicht sein.

Diese Hilfestellung richtet sich an Organisationen der Umweltbildung, die sowohl ihre **Wirkung evaluieren** wie auch ihre **Wirkung erhöhen** möchten. Die Hilfestellung unterstützt Sie, eine zweckmässige, effiziente Evaluation durchführen zu können, welche über eine Leistungserhebung und -berichterstattung hinausgeht. Über einfache Evaluationsfragestellungen hinaus wie «Wie viele Workshops haben wir durchgeführt?», «Wie viele Teilnehmende waren da?», «Wie gut hat der Workshop den Teilnehmenden gefallen?», möchten wir Ihnen Ideen geben, wie Sie «echte Wirkungen» wie Einstellungs-, Verhaltensänderung oder Kompetenzentwicklungen erheben können, und zwar auf schlanke und einfache Weise im Sinne einer Selbstevaluation. Durch ausgewählte Informationen, klare Handlungsempfehlungen und Best-Practice-Beispiele hoffen wir, Ihnen die Planung und Durchführung Ihrer eigenen Evaluation zu erleichtern und Sie zu ermutigen, eine Wirkungsevaluation mit Freude anzugehen.

Vorweg ist uns wichtig zu betonen, dass bereits die Auseinandersetzung mit einer Wirkungsevaluation vielfältige Lern- und Verbesserungsgelegenheiten für Sie bieten wird. Wenn Sie sich überlegen, welche Indikatoren Sie verwenden könnten, um Ihre Wirkungen sichtbar zu machen, denken Sie Ihr Angebot automatisch präziser und somit wirkungslogischer. Die gängigen Wirkungen, welche mit Umweltbildung und der Bildung für Nachhaltige Entwicklung verbunden werden – Kompetenzen, Einstellungen, Handlungen –, scheinen schwer zu fassen. Das liegt auch, aber nicht nur an methodischen Herausforderungen. Unterschätzt wird die Schwierigkeit, die Wirkungsziele eines Angebots konkret zu benennen: Welche direkten Wirkungen kann Ihr Umweltbildungsangebot haben? Wie lassen sie sich auf den Punkt bringen? Und in welcher Verbindung stehen diese Outcomes zu längerfristigen Veränderungen, dem Impact? Sie werden sehen, dass Sie neue Antworten auf solche Fragen erhalten, wenn Sie diese Hilfestellung durcharbeiten. Und natürlich erhalten Sie aus der durchgeführten Evaluation – auch wenn sie schlank gehalten ist – wichtige Anhaltspunkte, was funktioniert und wie Sie Ihr Angebot weiterentwickeln können.

«Keine Wirkungsevaluation ohne Wirkungsmodell» – dies ist das Leitmotiv dieses Leitfadens. Und entsprechend gestaltet sich sein **Aufbau**:

- ☞ Kapitel 1: Wirkungsmodell – Ohne Ziel kein Weg. Wie ist ein Wirkungsmodell aufgebaut? Wie lassen sich die direkten Wirkungen eines Angebots von den Leistungen und den indirekten Wirkungen abgrenzen?
- ☞ Kapitel 2: Outcomes – Wirkungen formulieren. Wie können die Outcomes eines Angebots präzise formuliert werden, sodass sie sich evaluieren lassen?
- ☞ Kapitel 3: Indikatoren – Wirkungen greifbar machen. Wie lassen sich Indikatoren bestimmen, um die Erreichung der Outcomes zu erfassen?
- ☞ Kapitel 4: Methoden – Wirkung messen. Wie gehe ich methodisch vor, um die Wirkungen zu messen? Welche Methoden gibt es über Befragungen hinaus?

Wirkungsmodell – Ohne Ziel kein Weg

Aus zwei Gründen ist es unabdingbar, die Wirkungsevaluation eines Angebots oder Projekts an dessen Wirkungsmodell auszurichten: Erstens lässt sich mit Hilfe eines Wirkungsmodells fokussiert und effizient evaluieren, da das Modell die beabsichtigten Wirkungen festhält. Das Wirkungsmodell gibt die Anhaltspunkte, welche Wirkungserwartungen in der Evaluation überprüft werden können. Ohne ein Wirkungsmodell als Ausgangs- und Orientierungspunkt für die Evaluation wüsste man nicht, nach welchen Wirkungen man eigentlich suchen muss. Eine Evaluation ohne Wirkungsmodell wird somit beliebig und ineffizient.

Zweitens bietet das Wirkungsmodell die Grundlage, um Wirkungen mit den Leistungen im Projekt in Verbindung zu setzen. Werden die im Wirkungsmodell formulierten wirkungslogischen Beziehungen in der Evaluation überprüft, werden Lernprozesse innerhalb der eigenen Institution möglich. Im Wirkungsmodell werden die Annahmen explizit gemacht, durch welche Leistungen die erwarteten Wirkungen eintreten sollen. Diese Annahmen werden in der Evaluation überprüft.

Aufgrund dieser Bedeutung, welche das Wirkungsmodell für eine Wirkungsevaluation hat, erläutern wir im Folgenden den Aufbau eines Wirkungsmodells.

1.1 Das Wirkungsmodell

Um die Wirkung und die Zusammenhänge von Projektzielen zu beschreiben, hat sich das IOOI-Modell als hilfreiches Werkzeug erwiesen. Das Modell umfasst vier Ebenen, die zum einen die Wirkungen eines Projektes (Impact und Outcome) und zum anderen den operationellen Leistungsbereich für die Umsetzung beschreiben (Input und Output).

Die Kernidee der Wirkungsorientierung ist es, dass ein Projekt von der Wirkung her geplant wird. Das, was im Projekt erreicht werden soll, legt die Basis für das, was im Projekt geleistet werden muss. Idealerweise wird somit ein Wirkungsmodell in der Reihenfolge «Impact – Outcome – Output – Input» erarbeitet.

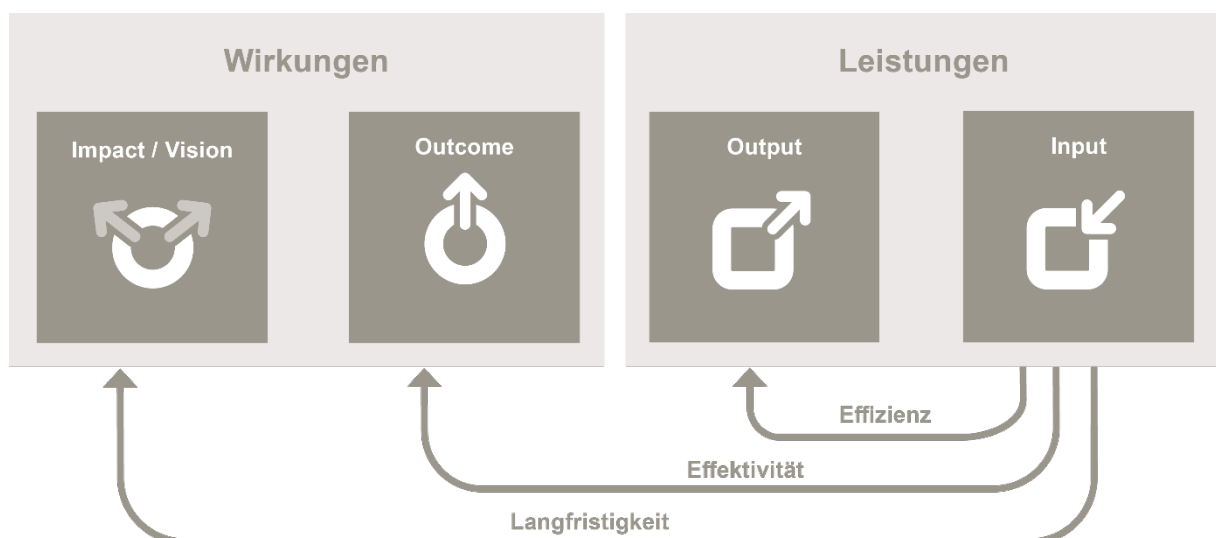


Abb. 1: Das IOOI-Wirkungsmodell mit den Wirkungsebenen Impact und Outcome und den Leistungsebenen Output und Input. Wirkungsorientierung bedeutet, dass in einem Projekt vom Impact her geplant und möglichst effizient, effektiv und langfristig umgesetzt wird (vgl. Müller & Wilhelm 2021).

1.2 Die Begrifflichkeiten

Impact / Vision



Beim Impact handelt es sich um die **indirekten, meist langfristigen Veränderungen**, zu denen ein Projekt beitragen soll, die aber nicht durch das Projekt allein bewirkt werden. Aus Sicht der Projektplanung stellt der Impact somit die Vision dar, welche positiven Veränderungen in der Gesellschaft oder der Umwelt eintreten sollen. So gesehen kann der Impact auch Veränderungen in der Zielgruppe umfassen, sofern bei der Projektplanung davon auszugehen ist, dass diese Veränderungen ausserhalb des direkten Einflussbereichs liegen. Häufig gilt dies für Verhaltensänderungen, die lange nach dem Besuch eines Bildungsangebots bei den Teilnehmenden eintreten sollen.

Beispiel:

Die Schweiz reduziert die Gesamtumweltbelastung durch Konsum auf ein Niveau, das für den ganzen Planeten vertretbar ist.

Die Schweizer Bevölkerung konsumiert ressourcenbewusst.

Outcome



Die Outcomes bezeichnen die **direkten Wirkungen** eines Projekts, welche zum Impact beitragen. Sie stellen in diesem Sinne das «Versprechen» dar, welche Wirkungen mit einem Projekt erreicht werden. Outcomes beziehen sich in der Regel auf Veränderungen bei den Zielgruppen eines Projekts. Sie werden immer aus der Sicht der Zielgruppe formuliert («Die Teilnehmenden können ...»). Die Evaluation des Erreichungsgrads der Outcome-Ziele liefert den Nachweis, ob das Versprechen eingehalten werden konnte. Outcome-Ziele werden so weit möglich mit konkreten Zielgrössen, Indikatoren und Erhebungsmethoden formuliert.

Beispiel:

Die Kursteilnehmenden analysieren ihren Ressourcenverbrauch und leiten 5 relevante Massnahmen her, wie sie suffizienter leben können.

Die Kursteilnehmenden nehmen sich vor, innerhalb der nächsten 2 Monate drei der Massnahmen umzusetzen.

Die Teilnehmenden reduzieren ihren Ressourcenverbrauch.

Output



Als Outputs werden diejenigen **Leistungen** verstanden, welche im Projektverlauf realisiert werden, um die Outcomes zu erreichen. Die Outputs weisen somit eine wirkungslogische Verbindung zu den Outcomes auf. Auf Output-Ebene werden auch quantitative Zielgrössen zur Anzahl der Angebote oder der Anzahl der Teilnehmenden festgelegt. Diese selbst sind noch keine Wirkungen, sie sind aber die Grundlage dafür, dass Wirkungen entstehen können.

Beispiel

Online-Kurs mit Übungen zum Thema «Ressourcenschonend konsumieren», 4 Durchführungen pro Jahr, je 30 Teilnehmende.

Input



Der Input eines Projekts umfasst alle **Ressourcen**, die zur Realisierung der geplanten Outputs notwendig sind. Dazu gehören finanzielle, personelle, aber auch immaterielle Ressourcen wie bereits vorhandene Kompetenzen oder Netzwerke.

Beispiel

Personalkosten für die Durchführung, Kooperation mit Konsummagazin für die Bewerbung.

Outcomes – Wirkungen formulieren

Die Outcomes eines Wirkungsmodells stellen den zentralen Gegenstand einer Wirkungsevaluation dar, da es sich bei ihnen um die direkten Wirkungen eines Projekts handelt. Die Formulierung griffiger, messbarer Outcomes bereitet in der Praxis oft Schwierigkeiten, weshalb wir näher darauf eingehen.

2.1 Ein «guter» Outcome...

Ein «guter» Outcome ist wie folgt formuliert. Der Outcome ...

- ... ist **aus Sicht der Zielpersonen und der bewirkten Veränderung** (des «Endverhaltens») formuliert. «Die Teilnehmenden lernen Vogelstimmen kennen», beschreibt, was im Kurs geboten wird. Ein besserer Outcome lautet: «Die Teilnehmenden können Vögel anhand ihres Gesangs bestimmen».
- ... nennt **aktive Handlungen oder sichtbare Veränderungen**, über die er beobachtet werden kann. Ein Outcome ist deshalb in spezifischen, sichtbaren und damit sprachlich präzisen und messbaren Verben beschrieben. Es gibt inspirierende Zusammenstellungen möglicher Verben, die einem die Formulierungen erleichtern (vgl. die Übersicht in Meyer 2020). Entsprechend ist die Formulierung «Die Teilnehmenden kennen 5 Vogelstimmen» zu ungenau. Besser ist: «Die Teilnehmenden können 5 Vogelstimmen zuordnen» oder «Sie können 5 Vogelstimmen imitieren».
- ... ist **SMART** formuliert, das heisst:
 - Specific** (spezifisch): Der gewünschte Outcome ist eindeutig beschrieben und definiert.
 - Measurable** (messbar): Die Zielerreichung ist messbar und kann überprüft werden.
 - Achievable / attainable** (erreichbar): Die Ziele sind realistisch und mit den verfügbaren Ressourcen erreichbar.
 - Relevant**: Die Ziele stehen im Einklang mit den übergeordneten Zielen und Strategien bspw. der eigenen Organisation oder einer Förderinstitution.
 - Time-bound** (terminiert): Die Ziele sind terminiert, also mit einer Zeitspanne versehen, bis wann sie erreicht werden sollen. Outcome-Ziele sind in der Regel – als direkte Wirkungen eines Projekts – innerhalb der Projektlaufzeit erreichbar.
- ... enthält eine angestrebte **Zielgrösse (einen Sollwert)**, welche auch via den Indikator definiert werden kann (vgl. Kap. 3.2 «Good Practice-Indikatoren»).

Die **Beispiele in Tabelle 1** zeigen, wie Outcomes formuliert sein könnten. Sie setzen sich aus einer ursprünglichen, ungenauen Form, dem Verbesserungspotenzial und möglichen Präzisierungen zusammen. Erst durch die Präzisierung wird eine fokussierte Wirkungsmessung möglich. Um ein breites Spektrum an möglichen Outcomes abzubilden, sind wir bei der Auswahl der Beispiele dem Handlungsmodell «Wissen – Können – Wollen – Tun» gefolgt (vgl. Müller & Wilhelm 2017):

- **Wissen**: Problembewusstsein, System- oder Sachwissen, Handlungs- oder Anwendungswissen, Wirksamkeits- und Selbstwirksamkeitswissen.
- **Können**: Kognitive, sensomotorische, soziale oder emotionale Fertigkeiten und Fähigkeiten, um Handlungssituationen erfolgreich zu bewältigen.
- **Wollen**: Bedürfnisse, Motive, Überzeugungen bis hin zu Absichten, um Handlungen auszuführen.
- **Tun**: Das tatsächliche Handeln. Erst im Tun wird eine Kompetenz (inklusive einer Handlungsabsicht) sichtbar.

In Bildungsangeboten zielen die direkten Wirkungen häufig auf Kompetenzen ab, welche die Teilnehmenden erwerben sollen. Unter Kompetenzen werden «Fähigkeiten, Fertigkeiten und Wissen, aber auch Bereitschaften, Haltungen und Einstellungen [verstanden], über die Lernende verfügen müssen, um neuen Anforderungen gewachsen zu sein» (Weinert 2001). Eine Kompetenz umfasst somit Aspekte des Wissens, des Könnens und des Wollens. Sie stellt die Bedingung dar, damit eine Handlung intentional ausgeführt werden kann. Zielt ein Projekt auf Verhaltensänderungen, sind entsprechend meist auch Outcomes auf den Ebenen Wissen, Können und Wollen notwendig, damit die Vorbedingungen der Verhaltensänderung gegeben sind.

2.2 Good Practice-Outcomes

Die folgende Tabelle 1 zeigt Beispiele von Outcomes, Aspekte mit Verbesserungspotential und mögliche Präzisierungen, wobei selbstverständlich weitere Varianten vorstellbar sind. Indem die Tabelle möglichst typische Beispiele von Umweltbildungs-Outcomes abbildet, sollten Sie nützliche Hinweise für Ihr Projekt wiederfinden können. Die **grau markierten Beispiele A-E** werden in den folgenden Kapiteln weiter vertieft.

Tabelle 1: Praxisbeispiele von Outcomes und mögliche Verbesserungsvorschläge.

Nr.	Ursprüngliche Version der Formulierung	Verbesserungspotenzial	Mögliche Präzisierung (wissen – können – wollen – tun)
A	«Schüler*innen verstehen den menschengemachten Klimawandel.»	- Unspezifische Beschreibung der Zielgruppe. - Unspezifisches Verb «verstehen»: Es wird nicht klar, wie sichtbar ist, ob etwas verstanden wird; «verstehen» ist nicht messbar.	«Schüler*innen der 8. Klasse können die Auswirkungen menschlicher Aktivitäten auf das Klima mithilfe des Indikators der CO2-Äquivalente beschreiben.» (wissen)
B	«Kursteilnehmende können Vögel richtig bestimmen.»	Eingrenzung der zu bestimmenden Vögel fehlt, genauso wie Art der Bestimmung.	«Teilnehmende können die 20 häufigsten einheimischen Brutvögel anhand ihres Aussehens, Rufs und Gesangs bestimmen.» (können)
C	«Teilnehmende sind sensibilisiert zum Thema Suffizienz.»	«Sensibilisierung» ist unspezifisch: Was bedeutet «sensibilisiert» sein? - Sofern die «Sensibilisierung» eine Folgewirkung bezweckt, kann diese erwähnt werden.	«TN können ihren Ressourcenverbrauch analysieren und Massnahmen herleiten, wie sie suffizienter leben könnten.» (können) «TN reduzieren ihren Ressourcenverbrauch.» (tun)
D	«TN verstehen eine offene Gesprächskultur und können für die Umwelt argumentieren.»	«Offene Gesprächskultur» ist zu unspezifisch - Unspezifisches Verb «verstehen»: Es wird nicht klar, wie sichtbar ist, ob etwas verstanden wird; «verstehen» ist nicht messbar. - Zwei Wirkungen in einem Outcome.	«TN versetzen sich in die Denkweise des Gegenübers.» (tun) «TN können ihre Sichtweise zu einem Umweltthema nachvollziehbar formulieren.» (können)
E	«TN wollen etwas gegen Foodwaste tun.»	Unspezifische Formulierung, Übersetzung in eine konkrete Handlung fehlt.	«Teilnehmende setzen Projekte zur Reduktion von Foodwaste in ihrem Umfeld um.» (tun)
f	«300 Klassen erleben den Biber in seinem natürlichen Lebensraum und reflektieren dabei ihre Beziehung zur Natur.»	«Die Beziehung zur Natur reflektieren» ist unspezifisch und wird erst messbar, wenn das Ziel konkreter formuliert ist. «300 Klassen, die an einer Biber-Exkursion teilnehmen», ist eine Zielgrösse auf Output-Ebene und formuliert noch keine Wirkung. - Zwei Wirkungen in einem Outcome; unspezifische Zielgruppe.	«Schüler*innen der 4. Klasse können die Lebensweise des Bibers in seinem natürlichen Lebensraum beschreiben.» (wissen) «Schüler*innen der 4. Klasse können in eigenen Worten die Bedeutung der Natur für sie persönlich formulieren.» (können)
g	«TN sind begeistert von der Vielfalt der Bodenfauna.»	«Begeisterung» ist unspezifisch und wird erst messbar, wenn das Ziel konkreter formuliert ist. «Begeisterung» beschreibt Wirkung während des Kurses. Es bleibt unklar, ob und wie die «Begeisterung» über den Kurs hinaus anhalten soll.	«TN können ihre Begeisterung (in eigenen Worten oder Bildern) ausdrücken.» (können) «TN möchten in ihrem Umfeld auch andere für die Vielfalt der Bodenfauna begeistern.» (wollen)
h	«Alle Leiter*innen nutzen Methoden der naturbezogenen Umweltbildung und reflektieren sie.»	- Unspezifische Zielgruppe «nutzen» allein sagt noch nichts über die Qualität des Gebrauchs der Methoden aus. - Reflexion ist als Überbegriff unspezifisch.	«Leiter*innen können mit dem Methodenpool umgehen, gezielt passende Methoden suchen und den Entscheid für eine Methode begründen.» (können) «Leiter*innen wenden mind. eine der vier Kernmethoden der naturbezogenen Umweltbildung regelmässig an.» (tun)

Indikatoren – Wirkungen greifbar machen

Nun, da Sie Formulierungsweisen für Outcomes kennengelernt haben, erfahren Sie in diesem Kapitel, wie man Outcomes messen kann. Um die direkten Wirkungen eines Angebots bestimmen zu können, wird in einem wirkungslogisch (also von der Wirkung ausgehend) geplanten Angebot die Erreichung der Outcome-Zielsetzungen gemessen. Manchmal gestaltet sich eine Outcome-Evaluierung relativ einfach. Handelt es sich beim Outcome um ein Wissensziel – beispielsweise, wenn die Lebensweise des Bibers beschrieben werden soll –, fällt es nicht schwer, sich eine Evaluation auszudenken. Wis-sensevaluationen sind uns aus dem schulischen Kontext vertraut und häufig gibt es ein klares «Richtig oder Falsch».

Schwieriger wird eine Evaluation, wenn die Teilnehmenden sich eine Fähigkeit, eine Einstellungsänderung, eine Motivation aneignen oder gar ihr Verhalten ändern sollen. In solchen Fällen ist häufig unklar, wie Sie die Wirkung erkennen können. Sie benötigen einen Indikator, der Ihnen die Wirkung anzeigt.

3.1 Indikatoren entwickeln

Das Wort «Indikator» bedeutet «auf etwas zeigen/etwas zeigen» (Beywl et al. 2017). Indikatoren sind mögliche Einheiten, die eine Zielerreichung erkennbar und messbar machen. Zählbare Indikatoren verwenden dabei einen numerischen Wert, z.B. die Anzahl richtig bestimmter Vogelarten. Da gerade motivationale oder emotionale Wirkungen nicht direkt zähl- oder sichtbar sind, ist eine Übersetzung in stellvertretende bzw. beschreibbare Indikatoren nötig. In der Praxis soll die vorgängige Beschreibung es so leicht wie möglich machen zu bestimmen, ob eine Zielwirkung erreicht wurde. Zum Beispiel könnte man das «Verständnis über den ökologischen Fussabdruck» von Schüler*innen auf vier Messniveaus vom Laien bis zur Expertin beschreiben (vgl. Tabelle 2). Diese kann man im Kurs beobachten und dokumentieren.

Indikatoren werden als Teil der Wirkungsorientierung definiert. Sie können je nach Expertise entweder vom Evaluationsteam oder unter Einbezug der Zielgruppe bestimmt werden. Inspirationen dazu können ähnliche Projekte oder die Literatur liefern. Oder Sie entwickeln eigene Indikatoren und Messniveaus. Für die Erarbeitung eigener Indikatoren empfiehlt sich folgendes Vorgehen (Kurz et al. 2013):

1. **Ideen sammeln:** definierte Wirkungsziele als Ausgangspunkt nehmen und gemeinsam überlegen, wie die Wirkungen sichtbar werden und wie sie gemessen werden könnten.
2. **Strukturieren und verfeinern der Ideen:** zusammengetragene Ideen aussortieren, zusammenfassen und ergänzen. Ein oder mehrere Indikatoren pro Wirkungsziel definieren.

Iterativer Prozess: Erfahrungsgemäss führt die Suche nach dem richtigen Indikator fast zwangsläufig zur Erkenntnis, dass der ursprünglich formulierte Outcome noch zu ungenau formuliert ist oder dass nicht alle Beteiligten die gleichen Vorstellungen haben, welche Wirkungen gemeint sind. Es lohnt sich, im Sinne eines iterativen Prozesses noch einmal auf den Schritt der Präzisierung der Wirkungsziele zurückzugehen und den Outcome zu schärfen. Nehmen wir beispielsweise an, die konkrete Wirkung ist, dass «TN sich in die Denkweise des Gegenübers versetzen.» Wäre der Indikator hier die «Anzahl Nachfragen beim Gegenüber» würde vielleicht klar, dass die Anzahl per se nicht zeigt, ob man zugehört hat und empathisch ist. Dementsprechend wird entweder die Wirkung (z.B. Interesse am Gegenüber zeigen und Indikator Anzahl Nachfragen) oder der Indikator (z.B. Offenheit für andere Denkweisen mit 4 Messniveaus) angepasst.

Um die Arbeit mit Indikatoren zu illustrieren, greift Tabelle 2 die Beispiele A-E aus Tabelle 1 auf. Neben dem Wirkungsziel zeigt der ursprünglich gewählte Indikator einen möglichen Ansatz, wie besagtes Ziel gemessen werden könnte. Die Tabelle zeigt Verbesserungspotenzial auf und macht Vorschläge, wie der Indikator präzisiert werden könnte.

3.2 Good Practice-Indikatoren

Tabelle 2: Beispiele von Indikatoren und möglichen Präzisierungen.

Nr.	Wirkungsziel basierend auf Handlungsmodell (wissen – können – wollen – tun)	Ursprünglich gewählter Indikator	Verbesserungspotenzial	Mögliche Präzisierung: verbesserter Indikator
A	«Schüler*innen der 8. Klasse können die Auswirkungen menschlicher Aktivitäten auf das Klima mithilfe des Indikators der CO ₂ -Äquivalente erklären.» (wissen)	TN können ihren eigenen CO ₂ -Ausstoss korrekt berechnen.	Indikator misst nicht das Wirkungsziel «wissen», sondern die Anwendung von CO ₂ -Äquivalenten auf sich selbst. Entweder Zielwirkung umformulieren oder Indikator anpassen.	Version angepasster Indikator: Korrektheit der Beschreibung der Auswirkungen menschlicher Aktivitäten auf das Klima mithilfe des Indikators der CO ₂ -Äquivalente Ggf. Messniveaus: 1. Keine Beschreibung möglich 2. Ansätze einer Beschreibung vorhanden, jedoch unvollständig/teils falsch 3. Beschreibung umfassend, jedoch eher oberflächlich 4. Beschreibung in Breite, Tiefe und Aktualität Zielgrösse: Mindestens 80% der TN erreichen Niveau 4
B	«Teilnehmende können die 20 häufigsten einheimischen Brutvögel anhand ihres Aussehens, Rufs und Gesangs bestimmen.» (können)	20 TN besuchen den Kurs zur Vogelbestimmung.	Quantitativer Indikator (20 TN), beschreibt das Angebot, nicht die Fähigkeiten der TN.	Indikator: Anzahl richtig bestimmter Vögel am Ende des Kurses Zielgrösse: 80% der TN können die gesichteten/gehörten Vögel korrekt bestimmen Indikator: korrekter Umgang mit Bestimmungshilfen Ggf. Messniveaus: 1. Orientierung mit Hilfsmitteln gelingt (auch mit Unterstützung) nicht 2. Umgang mit Hilfsmitteln funktioniert einigermaßen, mit Unterstützung 3. Selbstständige Orientierung mit Hilfsmitteln gelingt mehrheitlich 4. Selbstständige Bestimmung mit Hilfsmitteln gelingt fehlerfrei Zielgrösse: Mindestens 80% der TN sind mindestens auf Niveau 3
C	«TN können ihren Ressourcenverbrauch analysieren und Massnahmen herleiten, wie sie suffizienter Leben könnten.» (können)	Umfrage zu Handlungsfeldern für Suffizienz im eigenen Umfeld.	Methodenbeschreibung, anstatt Benennung der Einheit, die die Zielerreichung misst.	Indikator: Reflexionsfähigkeit Ressourcenverbrauch; Absichtserklärung zur Reduktion, Liste mit 3-5 Massnahmen Ggf. Messniveaus: 1. Eigener Verbrauch wird beschrieben, jedoch nicht hinterfragt 2. Eigener Verbrauch wird beschrieben, Struktur fehlt 3. Eigener Verbrauch wird beschrieben und mit dem Verbrauch anderer verglichen, Ansatzpunkte werden genannt 4. Ressourcenverbrauch wird beschrieben, Ansätze für Optimierung genannt 5. Absicht wird formuliert inklusive Bezug zum persönlichen Handlungsspielraum. Zielgrösse: Mindestens 60% der TN erreichen mindestens Niveau 3
D	«TN versetzen sich in die Denkweise des Gegenübers.» (tun)	Anzahl Nachfragen im Gespräch	Die Messbarkeit ist nur in Teilen gegeben, die tatsächliche Fähigkeit der TN (können) beschreibt der Indikator nicht.	Indikator: Grad des Perspektivenwechsels Ggf. Messniveaus: 1. TN stellen keine Fragen und bleiben bei ihrem Standpunkt 2. TN fragen aktiv nach und versuchen die Sicht des Gegenübers zu verstehen 3. TN fragen nach und versuchen, Lösungen mit dem/für das Gegenüber zu entwickeln. Zielgrösse: Mindestens 80% der TN erreichen mindestens Niveau 3
E	«Teilnehmende setzen Projekte zur Reduktion von Foodwaste in ihrem Umfeld um.» (tun)	Anzahl geplante Projekte zur Umsetzung.	Geplante Projekte (z.B. direkt nach dem Kurs) müssen nicht den tatsächlich umgesetzten entsprechen.	Indikator: Umgesetzte Projekte in der eigenen Organisation Zielgrösse: Mindestens 80% der TN setzen mind. 1 Projekt innerhalb von 6 Monaten nach Kursbesuch um

Methoden – Wirkungen messen

Nachdem die Indikatoren definiert sind, gilt es, die Methoden zu wählen, mit welchen die Wirkung über die Indikatoren gemessen werden soll.

4.1 Vorüberlegungen zur Methodenwahl

Folgende Überlegungen können helfen, die geeignete(n) Methode(n) zu finden.

- ☞ Die Methoden erlauben eine verhältnismässige Evaluation, das heisst, der Aufwand steht im Verhältnis zum **Zweck der Evaluation**. Wenn Sie verlässliche Wirksamkeitsnachweise (vgl. Kap. 4.2.1) möchten, scheuen Sie den dafür nötigen Aufwand nicht. Wenn Sie sich hingegen laufend verbessern wollen, kommen Sie mit wenig aufwändigen Methoden bereits weit (siehe unten).
- ☞ Die Methode ist an den **Gegenstand der Evaluation** angepasst. Bei Wirkungsevaluationen bedeutet dies, dass die Methode die Erreichung der Outcomes überprüfen kann. Die Methode kann die Veränderung durch ein Angebot abbilden und vom Zustand davor abgrenzen (bspw. Abgrenzung des Lerngewinns vom bereits vorhandenen Vorwissen).
- ☞ Wenn Sie eine Evaluation selbst umsetzen wollen (**interne Evaluation**), klären Sie ab, ob die Methode für Mitarbeitende ohne Zusatzausbildung und ohne kostenintensive Software verständlich und anwendbar ist.

Evaluationsmethoden können zu verschiedenen **Zeitpunkten** eingesetzt werden, je nach Zweck der Erhebung. Wollen Sie eine Veränderung feststellen, so erheben Sie ausgehend von den definierten Outcomes oder Kompetenzen, welche die Teilnehmenden erreichen sollen (siehe Kapitel 2 Outcomes – Wirkungen formulieren), in einer **Vor-Evaluation** den Stand vor dem Angebot. Soll beispielsweise ein Wissenszuwachs stattfinden, so interessiert das Vorwissen der Teilnehmenden. Eine solche Vor-Evaluation kann auch schlank umgesetzt werden, indem zu Beginn eines Angebots spielerisch der Wissensstand erhoben wird, bspw. durch ein Quiz oder über ein Abholen der Erwartungen der Teilnehmenden. Auch Einstellungen können vorab aktiviert werden, Bildkarten oder Metaphern sind dafür sehr geeignet. Werden nach dem Angebot die gleichen Methoden wieder verwendet, lässt sich die Differenz (bzw. die Wirkung) ermitteln, die durch das Angebot erreicht wurde (siehe Nach-Evaluation).

Nach der Vor-Evaluation kann das laufende Angebot **formativ** evaluiert werden. Der Grad der Outcome-Erreichung wird dabei fortlaufend überprüft, das Angebot gegebenenfalls justiert. Für diese formative Evaluation eignen sich Methoden, welche sowohl die Wirkung des Angebots unterstützen wie auch Hinweise auf die Outcome-Erreichung geben. Voraussetzung dafür ist ein flexibler Umgang mit der eigenen Planung während des Angebots.

In der **Nach-Evaluation** wird erhoben, ob die Outcomes am Ende des Angebots erreicht wurden. Im Verhältnis zur Vor-Messung ergibt sich dadurch die direkte Wirkung durch das Angebot, der klassische Vorher-Nachher-Vergleich. Die Nach-Evaluation kann beliebig um Messungen zu späteren Zeitpunkten ergänzt werden. Dadurch lässt sich die längerfristige Wirkung erheben (bspw. 6 Monate nach einem Angebot). Idealerweise werden zu allen Messpunkten die gleichen Methoden eingesetzt, damit die Evaluation aussagekräftig wird.

4.2 Auswahl an Methoden

Nachfolgend finden Sie eine Übersicht gängiger Evaluationsmethoden. Welche Methoden für Sie in Frage kommen, hängt – wie bereits erwähnt – vom Zweck Ihrer Evaluation, aber natürlich auch von den Ihnen zur Verfügung stehenden Ressourcen ab (Zeit, Finanzen, Personal etc.). Gerade im Rahmen einer Selbstevaluation lassen sich Methoden verwenden, die mit geringerem Aufwand umgesetzt werden können, aber dennoch zweckdienlich sind. Streben Sie aber einen eigentlichen Wirksamkeitsnachweis an, das heisst eine kausale Rückführung von Veränderungen auf Ihre Intervention, müssen Sie den Königsweg beschreiten, das Double-Difference-Vorgehen.

4.2.1 Der Königsweg der Wirkungsmessung

Damit Sie eine Wirkung eindeutig auf Ihre Intervention zurückführen können, müssen drei Bedingungen erfüllt sein: 1) Es muss ein systematischer Zusammenhang zwischen Intervention (Ursache) und Wirkung vorhanden sein, 2) die Ursache muss der Wirkung zeitlich vorausgehen und 3) andere alternative Einflussfaktoren müssen sich ausschliessen lassen. Gerade der letzte Punkt ist methodisch schwierig anzuwenden. Es könnte sein, dass Einflüsse von aussen für die Wirkung verantwortlich sind, die fälschlicherweise dem Angebot zugeschrieben werden. Beispielsweise könnte gleichzeitig zu einem mehrwöchigen Vogelstimmenkurs ein Hype um Vogelstimmen-Klingeltöne ausbrechen. Die halbe Bevölkerung spielt mit den Klingeltönen und lernt dabei 30 verschiedene Vogelstimmen zu unterscheiden. Ihre Kursteilnehmenden würden in der Schlussevaluation hohe Vogelstimmenkompetenzen aufweisen, diese aber unter Umständen nicht dank Ihrem Kurs, sondern dank den Klingeltönen erworben haben.

Um solche Ausseneinflüsse (möglichst) ausschliessen zu können, wird ein sogenanntes Double-Difference-Vorgehen angewendet: Zusätzlich zu den Veränderungen in Ihrer Zielgruppe evaluieren Sie auch jene in einer Kontrollgruppe. Zielgruppe und Kontrollgruppe sollten sich in allen Merkmalen ähnlich sein mit dem einzigen Unterschied, dass nur die Zielgruppe in den Genuss Ihres Angebots kommt. Sämtliche Befragungen führen Sie sowohl bei der Zielgruppe wie auch bei der Kontrollgruppe durch, also beispielsweise je eine Befragung vor dem Angebot und eine nach Abschluss des Angebots. Die Wirkung Ihres Angebots ergibt sich dann aus dem Unterschied zwischen den Wirkungen in Ihrer Ziel- und der Kontrollgruppe nach Ihrem Angebot, jeweils abzüglich der Ausgangslage vor der Massnahme. Im Falle eines Vogelstimmen-Klingelton-Hypes würde sich zeigen, dass die Vogelstimmen-Kompetenz auch in der Kontrollgruppe zugenommen hat. Das Double-Difference-Vorgehen ermöglicht somit, Wirkungen relativ verlässlich einer Intervention zuzuschreiben..

4.2.2 Die gängigen Evaluationsmethoden

Ein gängiger Weg, eine Wirkung zu erheben, ist jener der **Befragung**, sei es mit Hilfe eines Fragebogens, über Interviews oder durch spielerische Befragungsarten wie bspw. ein Quiz. Die Fragen bilden dabei die vorweg definierten Indikatoren ab. So kann erhoben werden, zu welchem Grad bei den Teilnehmenden eine Wirkung erreicht wurde. Wenn Sie bspw. die gleichen Wissensfragen vor und nach einer Intervention stellen, können Sie die Veränderung im Wissen herausfinden (mit den oben unter «Double-Difference» genannten Einschränkungen).

In die gleiche Richtung wie Befragungen gehen **Tests**. Getestet werden kann der Stand des Wissens, des Wollens, des Könnens oder des Tuns. Klassisches «Abfragen» bzw. ein Test zeigt bspw. an, ob die Teilnehmenden die Vogelstimmen den Vögeln richtig zuordnen können. Wird die «Prüfung» sowohl vor wie auch nach einem Kurs durchgeführt, zeigt sich der Unterschied in der Kompetenzentwicklung. Konnten vor einem Kurs 3 Vogelstimmen und danach 15 Vogelstimmen bestimmt werden, hat sich das Können in der Vogelbestimmung deutlich entwickelt.

Da Befragungen oft als «langweilig» angesehen werden, können Sie gewisse Zielerreichungen auch beobachten. Bei einer **Beobachtung** ist zentral, sich als Beobachtende*r an einem Beobachtungsbogen zu orientieren, damit die Ergebnisse von verschiedenen Durchführungen eines Angebots miteinander verglichen werden können und die Beobachter*innen einem gemeinsamen Leitfaden folgen (Weiss, 2023). Auch hier bilden die Beobachtungskriterien die vorweg definierten Indikatoren ab. Die

Methode der Beobachtung eignet sich sehr gut, um Interaktionsmuster, soziale Verhaltensweisen, Wissensaneignung und Einstellungsbildung zu verfolgen. Dies setzt natürlich voraus, dass der Bildungsinhalt im Rahmen eines Workshops, Unterrichts, einer Diskussion oder einer anderen Art der Veranstaltung auf eine Weise stattfindet, die eine Beobachtung auch tatsächlich zulässt bzw. ein Outcome mit einem Verb formuliert wurde, welches eine beobachtbare Tätigkeit nennt. Angewendet auf unser Beispiel der Vogelstimmen kann beobachtet werden, inwiefern sich die Teilnehmenden für die Vogelstimmen interessieren und wie sich der Lernprozess während des Kurses entwickelt.

Wenn Sie nicht die Möglichkeit haben, bspw. Wissen über einen Vorher-/Nachher-Vergleich zu testen, können Sie auch die Methode der **Selbstbeurteilung oder Selbstdeklaration** anwenden. Anstelle der Aussenbeurteilung der effektiven Veränderung kann nach der selbsteingeschätzten Veränderung gefragt werden, bspw. dem Wissenszuwachs. So können die Teilnehmenden am Ende eines Kurses gefragt werden, wie viele Vogelarten sie nun mehr bestimmen können als vor dem Kurs. Dieser Zugang der Selbstdeklaration hat den Vorteil, dass er deutlich weniger aufwändig ist. Anstelle einer Vor- und einer Nachbefragung wird nur eine Nachbefragung benötigt, um die Veränderung erkennen zu können. Da keine Kontrollmöglichkeit der Angaben der Teilnehmenden besteht, ist die Verlässlichkeit einer Selbstdeklaration eingeschränkt. Dennoch können Selbstdeklarationen zweckdienlich sein. Wenn die Evaluation primär Hinweise liefern soll, ob das Angebot funktioniert, macht diese durchaus Sinn. Den Ansatz der Selbstdeklaration können Sie auch zu einem späteren Zeitpunkt anwenden. Wenn Sie beispielsweise erfahren wollen, ob die Teilnehmenden das im Kurs erworbene Wissen auch in die Tat umgesetzt haben, können Sie diese einige Wochen später erneut kontaktieren und befragen. Oder Sie arbeiten bereits im Kurs mit Anreizen (Wettbewerb, Good-Practice-Blog etc.), um die Teilnehmenden zu motivieren, von sich aus über ihre Umsetzungen zu berichten. Ein Hilfsmittel zur Selbstdeklaration könnte beispielsweise ein Begleitjournal sein, welches die Teilnehmenden für fortlaufende Notizen verwenden, oder ein Fragebogen oder ein Gespräch am Ende des Bildungsangebots.

4.2.3 Methoden, die mehr sind als Evaluationsmethoden

Stehen für Wirkungsevaluationen nur geringe Ressourcen zur Verfügung, sind insbesondere Methoden interessant, welche sowohl Evaluationsmethode wie auch wirkungssteigernde Lern- oder Transfermethode sind. Sie tragen zum einen dazu bei, die Wirkung oder den Transfer eines Angebots zu verstärken, geben aber zusätzlich Hinweise, welche für die Wirkungsevaluation gebraucht werden können. Ein Beispiel sind Absichtserklärungen («Commitments», «Verpflichtungserklärungen»), welche Teilnehmende am Ende eines Kurses abgeben, bspw. bezogen auf ihre Absicht, das im Kurs Gelernte in der Praxis umzusetzen. Die Absichtserklärung ist ein psychologisches Interventionsinstrument, um die Wahrscheinlichkeit zu erhöhen, eine Handlung auch tatsächlich auszuführen. Gleichzeitig stellen die gesammelten Absichtserklärungen einen Indikator dar, wie viele der Teilnehmenden sich vornehmen, eine Handlung auszuführen. Zu erwähnen ist, dass diese Methode ihre Grenzen hat: Da die Outcomes hier nicht umfassend gemessen werden und meist nicht alle Teilnehmenden eines Angebots mitmachen, bekommt man nur ein Teilbild einer Evaluation.

4.2.4 Methodische Inspirationen

Die folgenden Methodenskizzen sollen Ihnen Anregungen geben, wie Sie Evaluationen kreativ umsetzen können. Interessant an den Methoden ist, dass sie in der Regel den vorhin beschriebenen doppelten Zweck erfüllen: Sie sind sowohl wirkungssteigernde Methode als auch Evaluationsmethode.

Zur Vertiefung der Methode finden Sie im Kurzbeschrieb zusätzlich die jeweilige Quellenangabe. Im anschließenden Kap. 4.3 zeigen wir, wie einzelne der Methoden an einem Beispielindikator angewendet werden können.

Appenzeller Wadenbeisser	
Kurzbeschrieb	TN formen am Ende des Kurses Tandems. Im Tandem wird definiert, welche Handlung bis zum Zeitpunkt X umgesetzt werden soll. Optional kann die Absicht im Kurs mündlich vorgestellt werden. Zum Zeitpunkt X treten die Tandems wieder miteinander in Kontakt, besprechen den Stand der Dinge und mögliche nächste Schritte. Die «Wadenbeisser» sind somit die Tandempartner*innen, die sich gegenseitig an die Aufgabe erinnern.
Evaluation	Zeitpunkt: Nach dem Kurs Vorgehen: Rückmeldung der Tandempartner*innen an die Kursleitung bereits im Kurs anfordern. Erinnerung nach X Monaten. Die gesammelten umgesetzten Aufgaben werden für die Evaluation den Wirkungszielen des Kurses gegenübergestellt. Ressourcen: Tandembildung wird als Reflexionsmoment in den Kursabschluss integriert. Zusätzliche Ressourcen nötig für die Nachfrage bei den TN nach X Monaten und die Auswertung der Rückmeldungen (pro TN bis zu 30 Minuten).
Quelle	Andreas Schubiger: Wie Transfer gelingt. hep Verlag, 2019
Beispielindikator	Formulierte / dokumentierte Transfer- bzw. Handlungsabsichten Umgesetzte Projekte in der eigenen Organisation

Bilderweg/Systembild	
Kurzbeschrieb	TN bekommen den Auftrag, schrittweise eine Bildstrecke respektive ein Systembild zu erstellen. Die Idee ist, dass verschiedene Elemente des Themas ähnlich einer Concept-Map oder einer Mindmap miteinander in Verbindung gebracht werden. Durch die Verknüpfungen zeigen die TN, ob sie die Zusammenhänge verstanden haben. Dabei kann die Entstehungsgeschichte eines Problems geschildert werden, z.B. Klimawandel (Dimension Wissen), oder es können mögliche Handlungsoptionen/Szenarien, also Angriffspunkte im System, aufgezeigt werden (Dimension Tun).
Evaluation	Zeitpunkt: Vor, während oder nach dem Kurs Vorgehen: Je nach gewähltem Zeitpunkt ist die Methode als eine Art Lernjournal (schrittweises Aufbauen des Bilderweges/Systembildes) oder als Test (z.B. als Abschluss eines Kurses, Vorher-Nachher-Vergleich etc.) anwendbar. Es können entweder verschiedene Versionen des Bildes oder nur die Endversion fotografisch festgehalten werden. Eine reine Evaluation des fotografisch festgehaltenen Endproduktes kann schwierig zu interpretieren sein. Es empfiehlt sich die Kombination mit einer Kurzpräsentation im Kurs. Ressourcen: Methode kann in den Kurs integriert werden. Je nach Zeitpunkt und Auslegung ist der Aufwand pro TN für die Evaluation auf max. 30 Min. einzuschätzen.
Quelle	Hochschule für Agrar- und Umweltpädagogik (Hrsg.): Methoden BAG 2022
Beispielindikator	Reflexionsfähigkeit des Ressourcenverbrauchs

Blitzlicht	
Kurzbeschreibung	TN fassen in einem Satz respektive wenigen Sätzen ihre Antwort auf eine bestimmte Frage zusammen, z.B. «Warum sollten wir uns (nicht) für die Perspektive des Gegenübers öffnen?». Alternativ können Satzanfänge als Hilfestellung gegeben werden, z.B. «Klimawandel bedeutet...». Um individuelle Antworten zu fördern und eine gegenseitige Beeinflussung zu vermeiden, empfiehlt es sich, die TN kurz individuell nachdenken zu lassen, bevor sie ihre Antwort im Plenum äussern.
Evaluation	Zeitpunkt: Vorher, während, nach dem Kurs Vorgehen: Das Blitzlicht ist eine zeiteffiziente Methode, um zu überprüfen, wie ein Konzept verstanden wird bzw. was der aktuelle Stand der Teilnehmenden bezüglich des Wirkungszieles ist. Dementsprechend kann es sowohl zu Beginn, während, aber auch nach einem Kurs eingesetzt werden. Am effizientesten scheint es, jeweils eine Tonaufnahme zu machen und direkt den Stand auf einer Messskala pro TN einzutragen. Wird das Blitzlicht mehrmals gemacht, kann eine Veränderung pro TN abgebildet und im Anschluss evaluiert werden. Ressourcen: Kann in den Kurs integriert werden. Um die Ablenkung der TN so gering wie möglich zu halten und möglichst viele Daten zu sammeln, ist eine Moderations- und eine Evaluationsperson zu empfehlen. Die anschliessende Auswertung der TN-Entwicklung sollte max. 2 h in Anspruch nehmen.
Quelle	Qualitäts- und UnterstützungsAgentur: Methodensammlung, 2023
Beispielindikator	Korrektheit der Beschreibung der Auswirkungen mithilfe des Indikators der CO2-Äquivalente

Konzeptentwicklung	
Kurzbeschreibung	Die Zustimmung zu Aussagen zum behandelten Thema wird zu verschiedenen Zeitpunkten von den Teilnehmenden eingeschätzt. Auf vorgelegte Aussagen können die TN mit einer Skala von «Stimme überhaupt nicht zu» bis «Stimme sehr zu» antworten. Über 2-3 Zeitpunkte sollte sich ein Verlauf des Konzeptverständnisses ergeben.
Evaluation	Zeitpunkt: Vor, während, nach dem Kurs Vorgehen: Durch die Messung der Einschätzung derselben Aussagen über mehrere Zeitpunkte kann man ähnlich wie beim Blitzlicht einen Lernprozess abbilden. Alternativ zum schriftlichen Beantworten der Fragen mit Checkboxes könnte der Konzeptwechsel auch mit einer räumlichen Moderation im Stil der Sendung «1, 2 oder 3» umgesetzt werden. Dafür müssten die Zeitpunkte und Aufstellungen jeweils fotografisch festgehalten werden. Im Anschluss könnten die Aufnahmen miteinander verglichen werden und das Wirkungsziel dementsprechend evaluiert werden. TN könnten sich sogar selbst auswerten, was Ressourcen des Evaluationsteams spart und einen zusätzlichen Lerneffekt ermöglicht. Ressourcen: Kann in den Kurs integriert werden. Für die anschliessende Auswertung rechnen wir mit ungefähr 1h pro TN.
Quelle	Grospietsch & Mayer, 2018 (nur als gedruckte Quelle verfügbar)
Beispielindikator	Anteil Teilnehmende, die die Aussagen/Konzepte richtig einschätzen.

Ping Pong Argumentation	
Kurzbeschreibung	Die Moderation fordert, dass sich die Teilnehmenden gemäss ihrer Meinung zu einer bestimmten Aussage aufstellen. In Zweierpaaren unterschiedlicher Meinungen wird diskutiert. Im Anschluss werden die Rollen gewechselt, z.B. um die Perspektive des Gegenübers zu vertreten oder um zu überprüfen, ob die Teilnehmenden zugehört haben.
Evaluation	Zeitpunkt: Während dem Kurs Vorgehen: Summativ am Ende des Kurses. Der/die Kursleiter*in geht bei jeder einzelnen Gruppe vorbei, hört für maximal 2 Minuten der Gesprächsdynamik zu und markiert die TN auf einer Beurteilungsskala (als Beispiel vgl. Nr. D in der folgenden Tabelle 3). Alle Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen. Ressourcen: Kann in den Kurs integriert werden. Für die anschliessende Auswertung rechnen wir mit max. 1h.
Quelle	Hochschule für Agrar- und Umweltpädagogik (Hrsg.): Methoden BAG 2022
Beispielindikator	Grad des Perspektivenwechsels

Lebendes Diagramm	
Kurzbeschreibung	Die Teilnehmenden werden aufgefordert, sich bezüglich ihrer Kompetenzen im Raum aufzustellen (z.B. Achse von «kann ich schon gut» vs. «kann ich noch gar nicht»).
Evaluation	Zeitpunkt: Vor, während, nach dem Kurs Vorgehen: Die verschiedenen Stadien der Selbsteinschätzung können fotografisch festgehalten werden. Alle Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen. Idealerweise lässt sich so ein Kompetenzzuwachs feststellen. Ressourcen: Kann in den Unterricht integriert werden. Die Auswertung der gesammelten Daten dauert max. 1h.
Quelle	Hochschule für Agrar- und Umweltpädagogik (Hrsg.): Methoden BAG 2022
Beispielindikator	Korrektur Umgang mit Bestimmungshilfen

Zertifikat für mich	
Kurzbeschreibung	TN schreiben sich selbst ein Zertifikat zu bestimmten, der Evaluation dienlichen Fragen, z.B. Wenn ich an die Bestimmung von Vogelarten denke... • Welche Fähigkeiten und Kompetenzen habe ich erlernt? • Welche Erkenntnisse hat mir der Workshop gebracht? Oder ganz konkret: • Diese Vogelarten kann ich bestimmen...
Evaluation	Zeitpunkt: Während, nach dem Angebot Vorgehen: Nachdem die TN ihr Zertifikat geschrieben haben, sammelt die Kursleitung die Dokumente ein. An die TN wird kommuniziert, dass die Zertifikate zum Zeitpunkt X nach dem Kurs verschickt werden, u.a. als Erinnerung an die erworbenen Kompetenzen. In der Zwischenzeit können Sie die Kompetenzen mit den Wirkungszielen vergleichen. Ressourcen: Die anschliessende Auswertung nimmt ca. 30 Min. pro TN in Anspruch
Quelle	Hochschule für Agrar- und Umweltpädagogik (Hrsg.): Methoden BAG 2022
Beispielindikator	Selbsteinschätzung erworbener Kompetenzen

4.3 Good Practice-Methoden

Tabelle 3 greift wiederum die Beispiele A-E aus den Tabellen 1 und 2 auf. Ergänzt sind Vorschläge von Messmethoden, Verbesserungspotenzial sowie mögliche Präzisierungen.

Tabelle 3: Methodenbeschriebe mit möglichen Präzisierungen.

Nr.	Indikator	Mögliche Messmethode	Verbesserungspotenzial	Präzisierung [Methode], Vorgehen & Begründung, Evaluation
A	Version angepasster Indikator: Korrektheit der Beschreibung der Auswirkungen menschlicher Aktivitäten auf das Klima mithilfe des Indikators der CO₂-Äquivalente Ggf. Messniveaus: 1. Keine Beschreibung möglich 2. Ansätze einer Beschreibung vorhanden, jedoch unvollständig/teils falsch 3. Beschreibung umfassend, jedoch eher oberflächlich 4. Beschreibung in Breite, Tiefe und Aktualität Zielgrösse: Mindestens 80% der TN erreichen Niveau 4	«Kurs wird während 3 Stunden durchgeführt»	Erhebungsmethode fehlt.	Blitzlicht (siehe Kap. 4.2.4) <i>In 2-3 Sätzen formulieren TN, wie sie die Klimaauswirkung menschlicher Aktivität mit CO₂-Äquivalenten beschreiben würden. TN überprüfen gegenseitig und lernen selbst zu formulieren, doppelter Lerneffekt</i> Formativ zu Beginn und am Ende des Kurses (Vorher-Nachher-Effekt). 1-2 Personen (Moderation, Dokumentation mit Bewertungsbogen mit 4er Skala; notfalls Audioaufnahme und Bewertung im Anschluss). Alle Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen.
B	Indikator: Anzahl richtig bestimmter Vögel am Ende des Kurses Zielgrösse: Mindestens 80% der TN können die gesichteten/gehörten Vögel korrekt bestimmen	«Erreichung des Ziels mittels Fragebogen überprüft»	Methode und Umsetzung unspezifisch.	Quiz anhand Tonaufnahmen und Bildmaterial <i>TN füllen am Ende des Kurses individuell ein Quiz aus. Dazu werden Bilder bzw. Tonaufnahmen der verschiedenen Arten gezeigt. Die TN erkennen selbst ihren Lerneffekt durch Vorher-nachher-Vergleich.</i> Formativ zu Beginn und am Ende des Kurses (Vorher-Nachher-Effekt); ggf. eine Applikation verwenden oder Test auf Papier. Alle Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen.
C	Indikator: Reflexionsfähigkeit Ressourcenverbrauch Ggf. Messniveaus: 1. Eigener Verbrauch wird beschrieben, jedoch nicht hinterfragt 2. Eigener Verbrauch wird beschrieben, Struktur fehlt 3. Eigener Verbrauch wird beschrieben und mit dem Verbrauch anderer verglichen, Ansatzpunkte werden genannt 4. Ressourcenverbrauch wird beschrieben, Ansätze für Optimierung genannt 5. Absicht wird formuliert inklusive Bezug zum persönlichen Handlungsspielraum.	«Abfragen/mündliche Rückmeldungen»	Methode und Umsetzung unspezifisch. Mündliche Rückmeldungen könnte Interviews, Fokusgruppe, schriftliche Befragung o.ä. heissen.	Bilderweg/Systembild (siehe Kap. 4.2.4) <i>TN bekommen am Anfang des Kurses den Auftrag, schrittweise ein Bild zu erstellen. Im Kurs werden verschiedene Aspekte des persönlichen Ressourcenverbrauchs thematisiert. Die Kursleitung animiert nach dem jeweiligen Themenblock dazu, den neuen Aspekt in das individuelle Ressourcenbild zu integrieren. Am Ende des Kurses werden die Bilder und mögliche Ansatzpunkte präsentiert. TN erarbeiten schrittweise den Wissenstransfer in eigenen Alltag. Hoher Grad der Personalisierung möglich.</i> Formativ während Kurs. 1-2 Personen (Moderation, Dokumentation mit Bewertungsbogen mit 5er Skala; notfalls Bewertung der Plakate). Alle

	Zielgrösse: 60% der TN erreichen Niveau 3			Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen.
D	Indikator: Grad des Perspektivenwechsels Ggf: Messniveaus: 1. TN stellen keine Fragen und bleiben bei ihrem Standpunkt 2. TN fragen aktiv nach und versuchen die Sicht des Gegenübers zu verstehen 3. TN fragen nach und versuchen, Lösungen mit dem/für das Gegenüber zu entwickeln. Zielgrösse: Mindestens 80% der TN erreichen Niveau 3	«Rollenspiel mit unterschiedlichen Perspektiven»	Methode im Rahmen des Angebots benannt, wie der Perspektivenwechsel gemessen wird, fehlt jedoch. Ebenso fehlen Begründung und genaue Anwendung.	Ping-Pong Argumentation (siehe Kap. 4.2.4) <i>TN werden aufgefordert, zu zweit ein Alltagsproblem der Nachhaltigkeit, z.B. Fleischkonsum, 2x10 Minuten zu diskutieren. Dabei sollen sie gegensätzliche Positionen vertreten. Nach der ersten Runde werden die Rollen gewechselt und sie vertreten die vorige Rolle des Gegenübers.</i> Summativ am Ende des Kurses. Kursleiter*in geht bei jeder einzelnen Gruppe vorbei, hört für maximal 2 Minuten der Gesprächsdynamik zu und markiert die TN auf der Skala von 1-3. Alle Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen.
E	Indikator: Umgesetzte Projekte in der eigenen Organisation Zielgrösse: Mindestens 80% der TN setzen mind. 1 Projekt innerhalb von 6 Monaten nach Kursbesuch um	«Selbsteinschätzung direkt nach dem Kurs»	Misst nur Versprechen, nicht tatsächlich umgesetzte Projekte.	Appenzeller Wadenbeisser (siehe Kap. 4.2.4) <i>TN bilden während dem Kurs sogenannte Umsetzungstandems. Diese halten am Ende des Kurses schriftlich fest, was sie planen, umzusetzen. Der «Wadenbeisser» bzw. Tandempartner hat die Aufgabe, nach 2 Monaten beim Gegenüber den Stand der geplanten Projekte abzufragen. Für den definitiven Erhalt des Kurszertifikats wird jedes Tandem aufgefordert, eine kurze Dokumentation der Projekte (1-2 Seiten) 3 Monate nach Kursbesuch an die Kursleitung zu schicken. Soziale Verpflichtung durch Absprache mit anderen Kursteilnehmenden kann im Vergleich zur Selbsteinschätzung mehr dazu anspornen, die Projekte umzusetzen.</i> Summativ 6 Monate nach Kursabschluss. Alle Ergebnisse werden im Anschluss ausgewertet und mit den Wirkungszielen abgeglichen.

Weiterführende Literatur

Zum Abschluss dieser Hilfestellung finden Sie eine Übersicht über weitere Hilfsmittel und Quellen, die online zugänglich sind und welche Sie auf Ihrem Weg zur selbstständigen Wirkungsevaluation unterstützen. Darüber hinaus regen Sie die weiterführenden Hilfestellungen vielleicht an, nach einer Phase der Selbstevaluation, eine vertiefte Evaluation mit externer Unterstützung durchzuführen. Dies trägt zu einer noch wirksameren Umsetzung von Projekten im Bereich der Umweltbildung bei, bedingt allerdings die dafür notwendigen finanziellen Ressourcen.

Tabelle 4: Weiterführende Quellen für Evaluationshilfestellungen.

Quelle / Titel	Inhalt
wb-web - Kompetenz für Erwachsenen- und Weiterbildner/-innen des DIE (Deutsches Institut für Erwachsenenbildung): <u>Formulierungshilfen für Lernziele - Methoden - Material - wb-web</u>	Hilfestellung, wie Verben messbar formuliert werden können. Beispiele für kognitive, affektive und motorische Kompetenzen
InForm – Deutsche Initiative für gesunde Ernährung und mehr Bewegung, <u>Leitfaden Evaluation</u>	Leitfaden für Evaluationen mit Fokus auf Selbstevaluation von Präventionsmassnahmen und Förderung der Gesundheit
Better Evaluation, <u>www.betterevaluation.org/methods-approaches</u>	Umfassende Sammlung von Evaluationsansätzen und -methoden der Global Evaluation Initiative, gut veranschaulicht, in Englisch
Leitfaden zur Wirkungsorientierung der Stiftung Mercator Schweiz « <u>Projekte mit Wirkung</u> »	Niederschwellige Einführung in die Arbeit mit und Online-Tool zur Erstellung von Wirkungslogiken
SEVAL – Schweizerische Evaluationsgesellschaft	Beratungsstelle für Evaluationen, Evaluationsstandards, diverse Fachpublikationen zum Thema Evaluation
BAFU / Interface, <u>Leitfaden zur Entwicklung von Wirkungsmodellen. Wirksame Interventionen planen, darstellen und messen</u>	Leitfaden insbesondere für Fachpersonen von Verwaltungen, um sie bei der Entwicklung und Gestaltung von Wirkungsmodellen zu unterstützen (mit Fokus auf Politikzyklus).

Quellen

- Beywl, W., Kehr, J., Mäder, S., & Niestroj, M. (2017). Evaluation Schritt für Schritt: Planung von Evaluationen. (3. unveränderte Auflage). Hiba.
- Grospietsch, F., & Mayer, J. (2018). Lernen mittels Konzeptwechsel in der Hochschuldidaktik. Lehrerbildung in vernetzten Lernumgebungen. In: M. Meier, K. Ziepprecht & J. Mayer (Hrsg.). Lehrerbildung in vernetzten Lernumgebungen, pp. 149-161. Münster: Waxmann.
- IN FORM (2017). Leitfaden Evaluation. Berlin: Bundesministerium für Ernährung und Landwirtschaft (BMEL).
- Kurz B. & Kubek D. (2017). Kursbuch Wirkung – 4. Auflage. Berlin: PHINEO gAG.
- Meyer, R. (2020). Lernziele formulieren. https://arbowis.ch/images/downloads/didaktik/Lernziele_formulieren.pdf, 17.4.2024.
- Müller, U. & Wilhelm, S. (2017). Umweltbildung mit Wirkung. Wädenswil: ZHAW.
- Müller, U. & Wilhelm, S. (2021). Umweltbildung mit Wirkung – Anleitung für die Eingabe von Gesuchen und Offerten. Bern: Bundesamt für Umwelt BAFU.
- Prodinger, M., & Winzheim, M. (2022). Methoden BAG - Bildung aktiv gestalten. Methoden für die Erwachsenenbildung. Hochschule für Agrar- und Umweltpädagogik.
- Qualitäts- und UnterstützungsAgentur. (2023). Methodensammlung. Schulentwicklung. <https://www.schulentwicklung.nrw.de/cms/methodensammlung/methodensammlung/index.html/liste.php>, 12.4.2024.
- Redman, A., Wiek, A., & Barth, M. (2021). Current practice of assessing students' sustainability competencies: A review of tools. Sustainability Science, 16, 117–135.
- Schubiger, A. (2019). Wie Transfer gelingt: Warum wir nicht immer tun, was wir wissen. hep verlag.
- Weinert, F. E. (2001). Concept of competence: A conceptual clarification. In: D. S. Rychen & L. H. Salganik (Hrsg.): Defining and selecting key competencies, pp. 45–65. Hogrefe & Huber Publishers.
- Weiss, A. (2023). Beobachtungen in der Evaluation von Wissenschaftskommunikation. Grundlagen und Praxis. In: Niemann, P., van den Bogaert, V., & Ziegler, R. (Hrsg.): Evaluationsmethoden der Wissenschaftskommunikation, pp. 135-153. Wiesbaden: Springer Fachmedien.