Preprint

Evaluation der Bildungs-/Schulungsaktivitäten bzw. des -programms und Wirkungsmessung

Marcus Flachmeyer

Abstract

Dieser Artikel bietet eine grundlegende Einführung in die wesentlichen Aspekte der Evaluierung von Bildungs- und Ausbildungsaktivitäten und -programmen in der Jugendarbeit. Zunächst werden die ethischen Herausforderungen einer Evaluierung und die Bedeutung der Klärung der Ziele erörtert. Der Artikel hebt die Vor- und Nachteile externer und interner Evaluierungen sowie die beiden Evaluierungsformen hervor: „summative“ und/oder „formative“ Evaluierung. Ein weiterer Schwerpunkt liegt auf der praktischen Planung einer Evaluierung in sieben Schritten. Schließlich werden vier Evaluationsmodelle vorgestellt, die in der non-formalen Bildung angewendet werden können (Kirkpatrick, Kaufman, Phillips, CIRO), und drei Modelle der Programmevaluation erwähnt, die gut zum Selbstverständnis und zur Mission von Jugendorganisationen passen (Stake, Cousins & Earl, Fetterman). Der Artikel betont die Notwendigkeit, Evaluation als integralen Bestandteil der Jugendarbeit zu verstehen, um die pädagogische Praxis kontinuierlich zu verbessern.

Flachmeyer, M. (2022). Evaluation der Bildungs-/Schulungsaktivitäten bzw. des -programms und Wirkungsmessung. In Implementierung des STREAM- Entrepreneurship-Ansatzes in der Arbeit mit Jugendlichen. Leitfaden für Fach– und Leitungskräfte in der Jugendarbeit (S. 40–59): Erasmus+ Projekt-Partnerschaft STREAMpreneur.

Evaluation der Bildungs-/Schulungsaktivitäten bzw. des -programms und Wirkungsmessung

Einleitung

Heute gehören nicht nur die Bedarfserhebung, die Planung/Organisation und die Durchführung von unterschiedlichen Bildungsangeboten zum Methodenspektrum der Jugendarbeit, sondern auch deren Evaluation und Verbesserung. In der Praxis steht gleichwohl das „Machen“ im Vordergrund, verstanden als die Planung, die Organisation und die Durchführung. Evaluation spielt eine eher untergeordnete Rolle, sie scheinen auf als Zufriedenheitsabfragen und sind da ein bewährtes Mittel, alle Jugendliche zu Wort kommen. Die Evaluation von Bildungsprogrammen, in der Regel veranlasst durch den Geldgeber oder/und den Träger, löst hingegen bei vielen haupt- und ehrenamtlichen Mitarbeiter/innen in der Jugendarbeit, aber auch bei den Jugendlichen selbst, nicht gerade Begeisterung aus. „Na, wenn es sein muss ...“ so lautet nicht selten die Reaktion.

Dieser Beitrag strebt nicht an, das definitive Rezept aufzuzeigen, mit dem Begeisterung für Evaluation in der Jugendarbeit hervorgerufen werden kann. Dies liegt in den Händen der haupt- und ehrenamtlichen Mitarbeiter/innen, die immer wieder aufs Neue entscheiden müssen, ob und auf welche Art und Weise die Wirkungen ihres pädagogischen Handelns, ihrer Bildungsaktivitäten und der Bildungsprogramme systematisch erfasst, dargestellt und kommuniziert werden. Dieser Beitrag liefert für solche Entscheidungen einige Orientierungen. Weit ausgreifend bespricht er zuerst ethische Herausforderungen einer Evaluation und die Klärung der Zielfunktionen, geht dann weiter auf die Frage nach externer oder interner Evaluation ein, bevor er dann die beiden Evaluationsformen „summativ“ und/oder „formativ“ bespricht. Die praktische Planung einer Evaluation steht anschließend im Fokus, ergänzt mit vier in der non-formalen Bildung anwendbaren Evaluationsmodellen und drei Programmevaluationsmodelle, die dem Selbstverständnis und dem Auftrag von Organisationen der Jugendarbeit sehr entgegenkommen. In diesen Ausführungen wird – soweit es der Platz erlaubt – auf die Umsetzungsmöglichkeiten nicht nur in Organisationen der Arbeit mit Jugendlichen eingegangen, sondern auch auf die MINT-Bildung.

Begriffsbestimmungen

Zunächst werden die Begriffe, die bereits in der Kapitelüberschrift gebraucht wurden, bestimmt, um eine gemeinsame Ausgangsbasis für die Lektüre dieses Kapitels zu schaffen. Diese Bestimmungen sind eher formaler Natur, sie werden im weiteren Verlauf des Kapitels konkretisiert.

Evaluation ist sowohl als Prozess als auch als Ergebnis zu verstehen, Scriven bezeichet Evaluation als „process of determing the merit or the worth or the significance of something; or the product of that process.“ (1981, S. 53). Er definiert die darin enthalten zentralen Begriffe: „Merit“ ist der „intrinsic value as opposed to extrinsic or system-based value/worth“ (S. 94), „Worth“ ist „the system value by contrast with intrinsic value“ (S. 167) und „Significance“ ist „the overall, synthesized conclusion of an evaluation, may relate to social or professional or intellectual significance.” (S. 145). Die Gegenstände von Evaluationen können vielfältig sein, u.a. Programme, Projekte, Dienstleistungen wie Trainings, Produkte, Personal, Organisationen, Daten, Theorien, Politiken.

Bildungsaktivitäten und Bildungsprogramme bewegen sich im gesamten Spektrum informeller und non-formaler Bildung. Dabei ist ein Programm als das größere und komplexere Format zu verstehen, nämlich als ein ganzes Bündel von zusammenhängenden Maßnahmen oder Aktivitäten mit einem bestimmten langfristigen Ziel, in der Regel top down konzipiert, entwickelt und umgesetzt. Das Joint Committee on Standards for Educational Evaluation (JCSEE) definiert ein Programm als “orchestrated initiative that dedicates resources and inputs to a series of activities intended to achieve specific process, product, services, output, and outcome goals” (Yarbrough, et al., 2011, p. 291). Bildungsaktivitäten sind demgegenüber das kleinere und weniger komplexe Format, sie können Teil von Bildungsprogrammen sein, müssen es aber auch nicht. Bildungsaktivitäten z.B. in einem Jugendzentrum haben oft einen eher singulären Charakter, sie reagieren kurzfristig auf die sichtbaren oder geäußerten Bedürfnisse von Jugendlichen, ihre strategische Dimension bleibt

oft unklar. Sie können aber auch eine klare strategische Dimension haben, z.B. wenn es um die Kompetenzen von Führungs- und Leitungskräften in Jugendverbänden geht.

Trainingsaktivitäten und Trainingsprogramme unterscheiden sich voneinander ebenfalls in den gerade geschilderten Merkmalen: Trainingsprogramme sind ein Bündel von zusammenhängenden Trainingsmaßnahmen oder -aktivitäten mit einem bestimmten längerfristigen Ziel (s.o.). Training ist dabei als eine spezifische Form non-formaler Bildung in der Jugendarbeit zu verstehen. Trainings, so eine Definition aus dem europäischen Diskussionskontext zur Jugendarbeit, zielen darauf ab, junge Menschen durch die Entwicklung von Wissen und Kompetenzen für das persönliche und (zunehmend) berufliche Leben zu stärken, wobei die Trainingsinhalte intrinsisch relevant und nützlich sind (Council of Europe, 2021). Auf Seiten des Anbieters werden für solche Trainings in der Regel die Lehrziele (und mögliche Lernziele) formuliert, der Trainingsinhalt wird methodisch-didaktisch aufbereitet, das Training wird zeitlich strukturiert und ein Trainer bereitgestellt.

Wirkungsmessung ist eine Variante von Evaluation, die seit jeher besonderen Stellen wert hat. „Interessant ist doch, was hinten rauskommt ...“ – und ganz besonders interessant sind die Wirkungen. Scriven definiert wie folgt „Impact Evaluation: An evaluation focussed on outcomes or pay-off rather than process delivery or implementation evaluation“ (1981, S. 74). Outcomes are usually the post-treatment effects, there are often effects dunng treatment.

Evaluation und Verantwortung/Ethik

Egal, ob es sich um die Evaluation eines einzelnen Trainings oder eines komplexen Bildungsprogramms handelt, es sollten sich diejenigen, die eine Evaluation veranlassen oder durchführen, bewusst sein, dass sie mit einer Evaluation auf ein Fachgebiet begeben, das durch umfangreiche theoretische Grundlagen und vielfältige praktischen Erfahrungen geprägt ist. Wie bereits in der grundlegenden Definition zu sehen ist, werden bei einer Evaluation immer auch Bewertungen vorgenommen (die wiederum selbstverständlich auch noch einmal bewertet werden können). Evaluationen sollten immer auch mit dem Bewusstsein aufgesetzt werden, dass Werturteile im Vergleich zu nur bloßen Beschreibungen oder Analysen ein noch höheres Ausmaß an Verantwortung für das Vorgehen wie auch für die Ergebnisse erfordern.

Vorbehalte gegenüber Evaluation sind auf alle Fälle sehr erst zunehmen, berühren sie doch im Kern grundlegende Fragen von Macht und von Ethik. Solche Fragen sind der Evaluatoren-Community nicht fremd, im Gegenteil. Standards der Evaluation sind u.a. ein Versuch, für alle Beteiligten mehr Sicherheit auf diesem anspruchsvollen Terrain zu verschaffen. Wegweisend hierbei ist das Joint Committee on Standards for Educational Evaluation (JCSEE), das erstmals 1981 den Grundkonsens der Evaluationsdisziplin mit Evaluationsstandards kodifizierte. Es liegen in der aktuellen Version des JCSEE-Handbuchs der Evaluationsstandards (Yarbrough, et al., 2011) dreißig Standards vor, gegliedert in fünf Dimensionen: Nützlichkeit, Durchführbarkeit, Korrektheit, Genauigkeit und Rechenschaftspflicht. Diese Standards zielen auf eine verantwortungsbewusste Evaluation pädagogischer Programme, Projekte und Materialien ab. Sie sind als „Leitplanken“ zu verstehen, die bei der Ausformulierung eines konkreten Evaluationsvorhaben die Richtung auf dem Weg zu einem verantwortungsbewussten Evaluationsvorhaben weisen können.

Diese fünf Dimensionen oder Kategorien guter Evaluationsqualität, geprägt von Sorgfalt und Fairness, enthalten dreißig Standards. Jeder der dreißig Standards wird in dem JCSEE-Handbuch so beschrieben, dass zuerst Sinn und Zweck dargelegt werden, es folgen dann sehr detaillierte Richtlinien zur Anwendung des Standards, eine Auflistung häufiger Fehler und ein oder mehrere Anschauungsbeispiele für die praktische Anwendung des Standards. Dieser Aufbau unterstreicht die Absicht, tatsächlich auf die Praxis der Evaluation einzuwirken und zu verantwortungsvollen und ethischen Evaluationsvorhaben zu führen.

Wird eine Evaluation auf dem Feld der Kinder- und Jugendarbeit geplant und durchgeführt, stellen sich die bei einer jeden Evaluation im Raum stehenden Fragen nach Macht und Ethik noch einmal in ganz besonderer Weise. Evaluation ist aufgefordert, Jugendliche in einer Art und Weise zu adressieren, die ihrem kognitiven und emotionalem Entwicklungsstand entspricht und ihrer spezifischen Schutzbedürftigkeit Rechnung trägt. Evaluation, verstanden als Praxisforschung oder angewandte Forschung, kann dabei von einem schon einige Jahre währenden Diskurs in der Forschungsgemeinde profitieren, z.B. von dem ERIC-Projekt. Das Projekt

„Ethical Research Involving Children“ (ERIC) versteht sich als Forum und Impulsgeber für die Forschungsgemeinschaft mit dem Ziel, dass die Rechte von Kindern und Jugendlichen auch in Forschungsprozesssen vollumfänglich respektiert werden. In den ERIC-Leitlinien werden Forscher aufgefordert, sich mit der Komplexität ethischer Fragen und ihren Werten, Einstellungen, Überzeugungen und Annahmen auseinanderzusetzen und die Art und Weise zu erkennen, in der diese die Entscheidungen, die sie im Forschungsprozess treffen, beeinflussen. (UNICEF: Office of Research - Innocenti ERIC, 2013).

Zielfunktionen von Evaluation

Es hat sich als sehr sinnvoll erwiesen, an den Beginn konzeptioneller Überlegungen die Frage zu stellen, wozu die Evaluation dienen soll. Ein einfaches Modell von Zielfunktionen, das Stockmann (2004) skizziert hat, kann dabei behilflich sein. Stockmann unterscheidet grundlegend vier Ziele: (1) die Gewinnung von Erkenntnissen; (2) die Ausübung von Kontrolle; (3) die Schaffung von Transparenz, um einen Dialog zu ermöglichen; und (4) die Dokumentation des Erfolgs (Legitimation).

Zielfunktion 1: Die Gewinnung von Erkenntnissen (Erkenntnisfunktion)
Evaluationen sollen Erkenntnisse liefern, die den Auftraggebern der Evaluation und den Zielgruppen des Programms (des Evaluationsgegenstands) interessieren, in irgendeiner Weise wertvoll sind und nützlich.

Zielfunktion 2: die Ausübung von Kontrolle (Kontrollfunktion)
Stockmann geht davon aus, dass bei der Evaluation zwar in erster Linie das Interesse im Vordergrund steht, Erkenntnisse zu gewinnen, er sieht aber auch das Interesse vor allem von Auftraggebern, Kontrolle auszuüben. Dies ist Organisationen, die mit öffentlichen Mitteln arbeiten, sehr vertraut, geht es doch bei Zwischen- und Abschlussberichten oder Monitoring-Gesprächen immer auch um die Frage, ob und wie der vorgelegte Plan tatsächlich realisiert wird, welche Schwierigkeiten es gibt und wie damit umgegangen wird. „D.h. mit jeder Evaluation ist direkt oder indirekt auch eine Form von Kontrolle verbunden.“ (Stockmann, 2004, S. 4).

Zielfunktion 3: die Schaffung von Transparenz, um einen Dialog zu ermöglichen (Lernfunktion)
Stockmann (2004) sieht in Evaluationen auch immer die Chance für Transparenz und Dialog, ein Anliegen, das ganz besonders für die Arbeit mit Kindern und Jugendlichen essenziell ist. Aber es gilt generell: Wenn die durch eine Evaluation gewonnenen Erkenntnisse transparent gemacht werden, ermöglichen sie einen Dialog zwischen verschiedenen Anspruchsgruppen wie den Geldgebern, der durchführenden Organisation, den Zielgruppen und sonstigen Beteiligten und Betroffenen. Damit bietet sich immer auch eine Basis für gemeinsames Lernen.

Zielfunktion 4: die Dokumentation des Erfolgs (Legitimationsfunktion)
Evaluationen zielen sehr häufig darauf ab, einen bestimmten Mitteleinsatz (input) und ein bestimmtes Vorgehen (process) zu legitimieren und mit dem Ergebnis (output/oucome/impact) in Beziehung zu setzen. Es liegt auf der Hand, dass dies nicht immer ganz trivial ist und komplexe Evaluationsfragestellungen auch sehr aufwendige Evaluationsdesigns erfordert (die wiederum Geld kosten). Die Ergebnisse können durchführende Organisationen (ebenso Geldgeber) zur Legitimation der eigenen Arbeit nutzen.

Stockmann (2004) geht auch darauf ein, dass Evaluationen sehr oft auch "taktische" Funktionen zugeschrieben werden (S. 4) Von taktischer Funktion wird dann gesprochen, wenn die Ergebnisse von Evaluationen nur dazu verwendet werden sollen, um bestimmte, bereits getroffene Entscheidungen zu legitimieren, oder weil es einfach sein muss oder chic ist. Er verweist in diesem Zusammenhang auf Pollitt, der schon 1998 eine „pathologische Seite“ beklagte: „Unfortunately, however, there is a pathological side to these developments. There are temptations for politicians to use evaluations as baubles or as bolsters—as decorative symbols of modernity or simply as reinforcements for courses of action they had already decided upon for other reasons“ (S. 223).

Interne und externe Evaluation

Insbesondere bei der Evaluation komplexerer Bildungsaktivitäten wie Bildungsprogrammen oder auch etwas größeren, projektförmigen Bildungsaktivitäten stellt sich in der Praxis oft sehr schnell die Frage, ob ein Evaluationsauftrag nach außen vergeben wird (externe Evaluation) oder die Evaluation mit eigenem Personal

geplant und umgesetzt werden kann/soll (interne Evaluation). Die Antwort darauf wird häufig von trivialen Faktoren bestimmt. Ursache der Entscheidung für eine interne Evaluation ist häufig schlichtweg das fehlende Geld für eine externe Evaluation. Ist Geld da, dann versprechen sich Auftraggeber einer externen Evaluation in der Regel mehr Distanz, einen umfassenderen Blick und eine größere Objektivität; zugleich wollen sie manchmal so auch einem eventuellen Eindruck vorbauen, an einer objektiven Evaluation von außen nicht wirklich interessiert zu sein.

Interne Evaluation

Bei der internen Selbstevaluation liegt die Konzipierung, Planung und Durchführung (und ggf. Meta-Evaluation) vollständig in Händen der eigenen Organisation. Dies hat eine Menge Vorteile. So ist davon auszugehen, dass der Zutritt zum zu evaluierenden Praxisfeld ohne Probleme möglich ist und die Evaluation methodisch und zeitmäßig dicht in die pädagogischen Prozesse eingepasst werden kann, was die Datenerhebung erleichtert. Konzeptionell kommt dazu, dass die Organisation selbst ein tiefes Verständnis des zu evaluierenden Gegenstands hat. Sie ist auf diesem Hintergrund in der Lage, die richtigen Fragen zu stellen, die erhobenen Daten einzuordnen und sie nah an ihrer eigenen Praxis zu deuten.

Ein möglicher Nachteil könnte die fehlende Distanz zum Gegenstand sein. Dieser mögliche Nachteil hat mindestens zwei Facetten, (1) eine forschungsparadigmatische und (2) eine professionelle Facette. Forschungsparadigmatisch wird häufig angeführt, dass mangelnde Distanz zum Gegenstand die Objektivität der Evaluation beeinträchtigt oder gar unmöglich macht. Hier ist zu überlegen, ob ein objektivistisches Verständnis von Evaluation/Forschung tatsächlich dem Gegenstand und dem Vorhaben angemessen ist, oder ob nicht andere Ansätze z.B. der rekonstruktiven Sozialforschung oder der Aktionsforschung besser passen. Die zweite Facette, die der Professionalität (der Organisation), ist ebenfalls zu klären. Wesentliches Merkmal von Professionalität ist Reflexivität, also die Fähigkeit zur Betrachtung und Einordnung eigenen Handelns und zum Denken in Alternativen; sie ist tatsächlich allererste Voraussetzung für eine interne Evaluation. Sie dürfte im Zuge der Professionalisierung der Arbeit mit jungen Menschen mehr und mehr gegeben sein.

Ein weiteres Merkmal der internen Evaluation, das möglicherweise als Nachteil gesehen wird, ist die damit verbundene qualifikatorische Anforderung an die Organisation. Interne Evaluation, insbesondere wenn es um die Evaluation von Bildungsprogrammen oder -projekten geht, setzt spezifisches, disziplinäres Wissen über Evaluation voraus. Dem ist aber entgegenzuhalten, dass manche Organisationen eine solche Fähigkeit mittlerweile vorhalten, in besonderen Organisationseinheiten oder auch auf Führungskräfte-Ebene. Die Fähigkeit zur Selbstevaluation (als Variante der internen Evaluation) auf der Ebene der pädagogischen Mitarbeiter/innen wurde in den letzten Jahren im Zuge der Diskussion um die Wirksamkeit pädagogischen Handelns in Organisationen der Jugendarbeit systematisch adressiert (u.a. Abele & Erdmann-Kutnevic, 2016; Council of Europe, 2018; Council of Europe Youth Work, 2021; König, 2007; SALTO Training and Cooperation Resource Centre, 2021; YouthLink Scotland, 2021).

Externe Evaluation

Bei einer externen Evaluation zumeist größerer, komplexer Bildungsprogramme begibt sich die Organisation in die Hände externer Spezialisten und erwartet zu Recht hochspezialisiertes deklaratorisches und prozedurales Wissen. Dies ist ein erwartbarer Vorteil externer Evaluation, zumal angenommen werden darf, dass das externe Institut und seine Evaluator/innen unabhängig sind und die Standards verantwortungsvoller Evaluation umsetzen. Erwartet der Geldgeber eine externe Evaluation oder mangelt es der auftraggebenden Organisation an eigener Fähigkeit, ist die externe Evaluation das Vorgehen der Wahl.

Sich für eine externe Evaluation entscheiden, das bedeutet aber nicht, dass sich der Auftraggeber zurücklehnen kann und nur noch bezahlt. Eine erfolgreiche externe Evaluation ist von der Mitwirkung des Auftraggebers abhängig. Ein gutes Evaluationsinstitut entwickelt mit dem Auftraggeber zusammen das Evaluationsdesign, plant mit ihm zusammen das Vorgehen und setzt es in enger Abstimmung und mit notwendiger Unterstützung der auftraggebenden Organisation um. Es erstellt - entsprechend der getroffenen Vereinbarung – Zwischenberichte und den Abschlussbericht und präsentiert ggf. die Ergebnisse. Die eigene Organisation wird also entlastet und zugleich methodisch durch das Evaluationsvorhaben geführt.

Gleichwohl müssen mögliche Risiken und Probleme externer Evaluation in den Blick genommen werden. So ist es nicht ungewöhnlich, dass eine externe Evaluation auf der Arbeitsebene nicht unbedingt willkommen ist und Abwehrreaktionen auslöst. Das beeinträchtigt unter Umständen nicht nur die Qualität der Daten, die Evaluation passt sich auch nicht so in die pädagogischen Prozesse ein, wie dies vielleicht wünschenswert ist. Dazu kommt noch das Risiko, dass die Daten aufgrund fehlenden Insiderwissens falsch interpretiert werden (abhängig vom Evaluationsdesign) und Schlüsse gezogen werden, die nicht nur falsch, sondern auch unbrauchbar sind. Diese möglichen Probleme sind in der Evaluationsdiskussion lange bekannt und waren u.a. Impuls für die „fourth generation evaluation“ (Guba & Lincoln, 1989), die die Anspruchsgruppen und insbesondere diejenigen, die „beforscht“ werden, in die Deutung, Einordnung und Bewertung mit einbezieht. Hinter diesem Konzept steht zum einen ein konstruktivistisches Erkenntnis- und Wissenschaftsverständnis, zum anderen die Einsicht, dass eine Evaluation nur dann sinnvoll ist, wenn die Anspruchsgruppen in ko-kreativer Weise an der Evaluation teilnehmen, sich deren Ergebnisse zu eigen machen und diese intrinsisch motiviert dann als Teil ihres professionellen Selbstverständnisses in der Praxis berücksichtigen.

Formative und summative Evaluation

Die Begriffe der formativen und summativen Evaluation wurden von Scriven (1967) geprägt, sie gehen auf einen Austausch mit Lee Cronbach zurück, der einen völlig anderen Blick auf die Zielfunktion von Evaluation einnahm als Scriven selbst. Cronbach (1963) vertrat die Ansicht, dass die Evaluation eines Bildungsprogramms besonders nützlich ist, wenn sie bereits während der Durchführung des Bildungsprogramms Hinweise zur Verbesserung gibt (und nicht erst am Ende). Seiner Ansicht nach ist ein solcher Nutzen für die Lehrer weitaus wertvoller als die nachträgliche Beurteilung des Bildungsprogramms. Scriven (1967) griff diese Argumentation im Kontext einer Diskussion zu Rollen und Zielen von Curriculumevaluation auf. So schätzte er es ebenfalls als sinnvoll ein, schon während der Entwicklungsarbeiten zu testen und die Erkenntnisse im Entwicklungsprozess zu verarbeiten. Diese Form der Evaluation bezeichnete er als formative Evaluation (S. 41). Eine Evaluation, die ein Curriculum (oder irgendein anderes Produkt) retrospektiv in der Gesamtschau bewertet, bezeichnete er als summative Evaluation.

Evaluationsmodelle und die Evaluationspraxis bedienen sich heute vielfach beider Formen der Evaluation. So werden im Verlauf des Bildungsprogramms, des Bildungsprojekts oder der Bildungsaktivität an mehreren sinnvoll ausgewählten Zeitpunkten Daten mit dem Ziel erhoben, den Umsetzungsprozess positiv zu beeinflussen. Die Daten werden aufbereitet und bei einem partizipativen Ansatz auch umfassend in den Umsetzungsprozess der Maßnahme wieder eingeführt. Ein solches partizipatives Vorgehen ist gerade in der Arbeit mit Kindern und Jugendlichen ein Muss (vgl. u.a. Hart, 2008), aber kein Selbstläufer. Die Ergebnisse müssen immer zielgruppengerecht aufbereitet und mit dem pädagogischen Hauptprozess sinnvoll verbunden werden, ein Ansinnen, das in der Praxis nicht immer ganz leicht umsetzbar ist und sorgfältig geplant werden muss. Ist das Bildungsprogramms, das Bildungsprojekts oder die Bildungsaktivität dann abgeschlossen, werden erneut Daten erhoben, direkt nach Ende und/oder auch mit einiger Zeit Abstand. Sie sind Grundlage für die summative Evaluation, mit der die Maßnahme insgesamt in der Rückschau bewertet wird; Erkenntnisse der formativen Evaluation fließen dabei in der Regel mit ein.

Planung einer Evaluation

Die Anzahl der Anleitungen und Checklisten, die bei der Planung und Durchführung eines Evaluationsvorhabens behilflich sein können, ist inzwischen nahezu unüberschaubar. Dies spiegelt zum einen wider, wie sehr Evaluation längst über das über Jahre exklusive Gebiet der Bildungsprogramme hinaus in andere Gebiete wie Gesundheit, Umwelt, Verkehr, Gesellschaft u.a.m. hineingetreten ist. Zum anderen zeigt sich, dass sich Evaluation in der Art und Weise, wie sie geplant und umgesetzt wird, von innen heraus weiterentwickelt und ausdifferenziert hat, aber auch externen Trends unterliegt, wie sie z.B. im Projektmanagement sichtbar werden. Waren in den 1980er und 1990er Jahren noch sorgfältige und langwährende Analysen, Definitionen und Planungen und anschließend sehr linear eine planmäßige Implementierung des Projektplans der „state of art“, so sind heute Flexibilität und Agilität angesagt (Michaelides & Antonacopoulou, 2021). Insgesamt betrachtet gibt es also nicht nur eine unüberschaubare Anzahl, sondern auch eine Vielfalt von Herangehensweisen und nicht den einen Königsweg.

Welchen Weg eine Organisation in der Jugendarbeit und mit ihr die pädagogischen Mitarbeiter/innen beschreiten, ist auf diesem Hintergrund auch und ganz besonders eine Frage der Organisationskultur, der eigenen Fähigkeiten und der Praktiken der Organisation. Evaluation muss „passen“ und sich einbetten, damit sie die wesentliche Funktion, nämlich die eigene pädagogische Praxis zu verbessern, erfüllen kann. Insofern sind die folgenden Planungsschritte als Empfehlungen zu verstehen, die immer der Anpassung an die eigenen Bedürfnisse bedürfen, ohne allerdings die Evaluationsqualität zu beeinträchtigen.

Schritt 1: Anlass, Ziele und Gegenstand der Evaluation klären, Konsens herstellen

Es ist hilfreich, wenn direkt zu Beginn bei den unterschiedlichen Anspruchsgruppen und Nutzern Klarheit darüber besteht, um was es eigentlich bei der Evaluation geht. Diese Dringlichkeit besteht nicht nur bei der Evaluation von Bildungsprogrammen und Bildungsprojekten, sondern – entsprechend runtergebrochen – auch bei der Evaluation eines Trainings. Auch bei einer vergleichsweisen kleinen Einheit bedarf es Legitimation und Transparenz. Das fängt bereits beim Anlass an. „Wer ist denn auf diese Idee gekommen?“ – „Das ist Bedingung des Förderprogramms“, so kann dann z.B. die klare Antwort bei einem Bildungsprojekt lauten. Meistens ergibt sich aus dem Anlass bereits die Zielfunktion (s.o.), die mit dieser Evaluation verfolgt wird. „Was ist denn eigentlich die Stoßrichtung?“ Wenn es um Kontrolle geht, ist es hilfreich, wenn das auch ohne Umschweife deutlich gemacht wird. Vielleicht sind aber auch Schwierigkeiten aufgetaucht, deren Ursache dem federführenden pädagogischen Mitarbeiter nicht ganz klar sind, dann steht die Erkenntnisfunktion im Vordergrund. Und es sollte direkt auch der Gegenstand der Evaluation eingegrenzt werden. „Was soll in den Blick genommen werden?“ Diese Fokussierung hilft dabei, gut vorbereitet den nächsten Schritt zu tun.

Schritt 2: Kapazitäten und Ressourcen klären

Ist so konsensuell umrissen, um was es bei der Evaluation gehen soll, lässt sich schnell klären, ob in der Organisation überhaupt die erforderlichen Kapazitäten und Ressourcen für ein Evaluationsvorhaben zur Verfügung stehen. Dabei spielt es kaum eine Rolle, ob es um die Evaluation eines Bildungsprogramms, eines Bildungsprojekts oder eine einzelne Bildungsmaßnahme wie ein Training geht. Viel entscheidender für die Beurteilung der eigenen Kapazitäten und Ressourcen ist eine Vorstellung, um was es bei der Evaluation gehen soll (und welcher methodische und instrumentelle Aufwand damit vermutlich erforderlich ist).

Hat die Organisation den Eindruck, dass sie selbst über die fachlichen Kapazitäten für die Planung und Durchführung einer Evaluation verfügt, ist sie auch in der Lage, die erforderlichen internen Ressourcen zu beziffern. Es sollte an dieser Stelle ein erstes Budget umrissen werden, auf dessen Grundlage eingeschätzt werden kann, ob eine interne Evaluation in einem positiven Verhältnis zu den zu erwartenden Ergebnissen steht. Geht es um die Evaluation eines fremdfinanzierten Bildungsprojekts oder einer Bildungsmaßnahme, lassen sich die errechneten Ausgaben möglicherweise dort finanzieren.

Hat die Organisation den Eindruck, dass sie selbst nicht über die erforderlichen fachlichen Kapazitäten verfügt oder präferiert sie aus anderen Gründen eine externe Evaluation, sollte sie sich ein Angebot von einem fachlich versierten Institut unterbreiten lassen. Ein gut gemachtes Angebot wird nicht nur eine erste Auskunft über das Vorgehen des Instituts geben, sondern auch Auskunft zum Honorar und in etwa zu Art und Umfang der erforderlichen Mitwirkung.

Ist dieser Schritt getan, hat die Organisation ausreichend Klarheit hinsichtlich ihrer eigenen Kapazitäten und Ressourcen. In diesem Zusammenhang wurde auch die Frage einer internen oder externen Evaluation geklärt. Nun sollten die vorliegenden Informationen für eine Entscheidung für oder gegen eine Evaluation ausreichend sein.

Schritt 3: Arbeitsfähigkeit aufbauen

Ist diese Entscheidung getroffen, stehen umfangreiche Arbeiten an, für die personelle und möglicherweise auch materielle Ressourcen zur Verfügung gestellt werden müssen. Je nach Umfang des Evaluationsvorhabens handelt es sich vielleicht nur um wenige Arbeitstage einer einzigen Person, vielleicht nimmt aber nun auch ein Projektteam die Arbeit auf, welches zu der Organisation in Beziehung gesetzt werden muss. Handelt es sich um eine externe Evaluation, ist die Beziehung zwischen der eigenen Organisation als Auftraggeber und dem

externen Institut als Auftragnehmer vertraglich zu fixieren. Aber auch im Falle einer internen Evaluation braucht die einzelne Person oder das Projektteam eine eindeutige Legitimation und klare Spielregeln seitens der Organisation. Der interne Auftraggeber muss eindeutig bestimmt sein, die Beziehung zu einer denkbaren Lenkungsgruppe oder einem Beirat klar geregelt.

Schritt 4: Evaluationskonzept aufsetzen

Es ist zwar in der Praxis von Aufrufen und Ausschreibungen noch üblich, dass die Entwicklung eines Evaluationskonzepts, die Instrumentenentwicklung und die Durchführungsplanung noch vor der Bewilligung oder Zusage geschieht, das ist aber weder ressourcenschonend noch effizient. Die Entscheidung zur Planung und Durchführung einer Evaluation sollte jetzt bereits gefällt sein, um auch tatsächlich über die Ressourcen für die Erstellung eines Evaluationskonzepts in enger Abstimmung mit Anspruchsgruppen und Beteiligten zu verfügen.

Was aber ist unter einem Evaluationskonzept eigentlich zu verstehen. Im deutschsprachigen Glossar der Evaluation wird ein Evaluationsplan (alternativ: Evaluationskonzept) definiert als:

> Detaillierter, konsistenter und schriftlich festgehaltener Entwurf des beabsichtigten Vorgehens einer konkreten Evaluation. Er stellt den gewählten Evaluationsansatz dar, die vorzunehmenden Schritte und einzusetzenden Evaluationsmethoden (u. a. Datenerhebungsplan). Er benennt Zuständigkeiten, konkretisiert Verfahren der Berichterstattung sowie die beabsichtigte Nutzung der Evaluation. Die Planung richtet sich an dem vereinbarten Evaluationszweck und den Evaluationsfragestellungen aus. (Eval-Wiki: Glossar der Evaluation, 05.11.2020).

Mit Blick auf das spezielle Anwendungsfeld, die Arbeit mit Jugendlichen, sollte das Evaluationskonzept zusätzlich dazu ausführen, in welcher Art und Weise dem kognitiven und emotionalem Entwicklungsstand von Jugendlichen und ihrer spezifischen Schutzbedürftigkeit Rechnung getragen wird.

Ist dieser Schritt getan, liegt für alle Beteiligten und Anspruchsgruppen ein Rahmenwerk vor, das Auskunft zu den methodischen Entscheidungen und zum Vorgehen gibt.

Schritt 5: Instrumente entwickeln/zusammenstellen

Ist der Rahmen gesetzt, können die darin genannten Methoden instrumentiert werden. Dabei geht es zuerst um die Datenerhebung, dann um die Datenaufbereitung/-darstellung und Datenauswertung. Dafür steht grundsätzlich die ganze Palette der empirischen Sozialforschung zur Verfügung und damit auch das breite Feld quantitativer, qualitativer und Mixed-Method-Forschung.

Die Auswahl oder die Entwicklung der Erhebungsinstrumente setzt voraus, dass die Evaluationsfragestellungen auf die Ebene von Forschungsfragen heruntergebrochen sind. Forschungsfragen sind viel spezifischer und können den Forschungsprozess fokussieren. Sie beinhalten insbesondere bei einer quantitativen Wirkungsmessung idealerweise bereits eingeführte Konstrukte wie Selbstwirksamkeitserwartungen, Risikobereitschaft, Bedürfnis nach Erkenntnis u.a.m. und sind damit der Zugang zu bereits validierten Item-Batterien. Alternativ müssen Indikatoren und Items bei einem quantitativen Forschungsansatz erst noch entwickelt werden.

Sind die Forschungsfragen klar, können Fragebögen, Beobachtungsbögen, Interviewleitfäden, Gesprächsleitfäden etc. ausgewählt oder entwickelt werden. Werden vorhandene Instrumente genutzt, sind zuvor die Rechte zu klären. Die Instrumente müssen den spezifischen Validitätskriterien genügen. Die Instrumente als auch das Medium müssen den Befragten mit ihren kognitiven, motivationalen und kulturellen Merkmalen Rechnung tragen.

Die Verarbeitung der Daten, insbesondere qualitativer Daten, erfordert einen erheblichen Zeiteinsatz und eine angemessene Auswahl von technischen Hilfsmitteln. Quantitative Daten lassen sich auch mit MS-Excel erstaunlich weitreichend verarbeiten, eine qualitative Inhaltsanalyse der Ergebnisse einer Gruppendiskussion, eines narrativen Interviews, eines Soziodramas u.a.m. lässt im Rahmen eines kleinen bis mittleren Evaluationsvorhabens auch ohne Analysesoftware wie MAXQDA erledigen. Sind die Daten verarbeitet, ist bei

der Darstellung der erhobenen Daten daran zu denken, sie so aufzubereiten, dass sie Interesse wecken, verständlich sind und als nützlich erlebt werden. Diese zielgruppenspezifische Darstellung kostet Zeit (und damit Geld).

Die Auswertung und Bewertung der Daten sollte so geplant werden, dass die Anspruchsgruppen und Beteiligten durch die Evaluatoren/innen ermutigt werden, ihre Deutungen und Einordnungen einzubringen. Dies kann in Workshops passieren, Einzelgesprächen, beim Ergebnisbericht über die pdf-Kommentarfunktion u.a.m. Es sollte aber nicht unterschätzt werden, dass diese partizipative Bewertung stimulierende Formate und Ressourcen erfordert.

Schritt 6: Evaluation durchführen

Die Implementierung des Evaluationskonzepts und seiner Instrumente und Verfahren braucht vor allem eins: Vertrauen, Transparenz und Commitment. Wichtige Promotoren sind die Mitarbeiter/-innen, die diesem Evaluationsvorhaben uneingeschränkt Bedeutung beimessen, es glaubwürdig kommunizieren und so als Vorbild agieren. Förderlich ist auch, wenn Jugendliche, die als Schlüsselpersonen hervortreten, sich dafür stark machen, z.B. in den Sozialen Medien. Je nach Größe des Evaluationsvorhabens und seiner Bedeutung für die Organisation ist der Kommunikation über verschiedene Medien hohe Bedeutung beizumessen.

Schritt 7: Ergebnisse sichern und verwerten

Bereits im Evaluationskonzept wurden Verfahren der Berichterstattung sowie die beabsichtigte Nutzung der Evaluation konkretisiert, das schließt die finale Berichterstattung und die Nutzung der Endergebnisse mit ein. In einem guten Evaluationskonzept sind diese Maßnahmen so ausgestaltet, dass sie die Zielfunktion/en des Evaluationsvorhaben effektiv und effizient unterstützen. Es bleibt aber ein gewisses Maß an Unsicherheit, ob diese Maßnahmen so erfolgreich umgesetzt werden können wie dies angestrebt ist.

Eine strukturelle Herausforderung ist die zeitliche Lücke zwischen dem Ende der Bildungsaktivität und der Aufbereitung der Ergebnisse der summativen Evaluation. Die Aufbereitung der Ergebnisse in einem Bericht, einem Podcast oder Videocast erfordert in der Praxis eine gewisse Zeit, partizipative Formate wie ein Ergebnisworkshop für Anspruchsgruppen und Beteiligte sind mitunter schwer zu realisieren, weil die Jugendlichen nicht mehr ansprechbar sind und sich schon auf andere Dinge fokussieren.

Sollen die Anspruchsgruppen und Beteiligten die Ergebnisse tatsächlich zur Kenntnis nehmen, ist ihre Bindung an die Organisation und an die Bildungsaktivität/das Bildungsprojekt Voraussetzung. Wirksam sind auch Evaluationsansätze mit einem maximalen Mitwirkungsgrad wie z.B. die Empowerment-Evaluation (Fetterman, 1994). Sie begünstigen nicht nur die Kenntnisnahme der Ergebnisse, sondern auch die Verwertung im Sinne einer Verbesserung des Produkts.

Anwendung bestehender Evaluationsmodelle auf non-formale Bildung

Bei der Konzipierung des eigenen Evaluationsvorhaben, und sei es noch so klein, ist es sehr hilfreich, auf bewährte Evaluationsmodelle zurückzugreifen. Wenn hier von Modellen die Rede ist, so handelt es sich nicht um mathematische Modelle, die zur Aufstellung von Theorien verwendet werden, sondern um Konzepte und Strukturen, mit dem sich die jeweiligen Autoren dem zu evaluierenden Gegenstand nähern und die vielfach in der Praxis angewendet werden.

Die in diesem Abschnitt angeführten Modelle stammen allesamt aus der Trainingsevaluation. Auch wenn sie zum Teil schon etwas betagter sind und ihre Anfänge bis in 1950er Jahre zurück gehen, so sind sie gleichwohl heute noch aktuell. Insbesondere das 4-Ebenen-Modell von Kirkpatrick (zuerst: 1959a) ist heute in der Trainingsevaluation noch allgegenwärtig.

Modelle aus der Trainingsevaluation sind immer dann interessant für die Konzipierung eines eigenen Evaluationsvorhabens, wenn sich das zu evaluierende Produkt durch Intentionalität auszeichnet und Lehr/Lernziele vielleicht sogar regelrecht mit Hilfe von Lernzieltaxonomien (u.a. Anderson, et al., 2009; Bloom & Krathwohl, 1956; Krathwohl, et al., 1966; Simpson, 1966, 1972) aufbereitet sind. Dies ist in der außerschulischen Bildungsarbeit mit Jugendlichen teilweise der Fall. Die hier vorgestellte kleine Auswahl von

Trainingsevaluationsmodellen, und zwar die Modelle von Kirkpatrick (1959a, 1959b, 1960a, 1960b), Kaufman und Keller (1964), Phillips (1991) und Warr, Bird und Rackham (1970) weisen viele Gemeinsamkeiten, aber eben doch markante Unterschiede auf, die inspirierend für das eigene Evaluationsvorhaben sein können.

Das Kirkpatrick-Modell

International das vermutlich bekannteste Modell der Trainingsevaluation ist das Modell von Kirkpatrick. Kirckpatrick hat sein Modell für die Evaluation von Trainings der amerikanischen Hilfsorganisation Heifer International Ende der 1950er Jahre entwickelt 1959 und 1960 mit einer Serie von vier Artikeln erstmals publiziert wurde. Kirkpatrick strukturiert Trainingsevaluation mit vier Schritten (Kirkpatrick, 1959a, 1959b, 1960a, 1960b) bzw. Ebenen (Kirkpatrick & Kirkpatrick, 2006): (1) Reaktion, (2) Lernen, (3) Verhalten, und (4) Ergebnisse.

Auf der Ebene **(1) Reaktion** wird die unmittelbare Reaktion der Teilnehmer/innen evaluiert, ausgedrückt am ehesten in Begriffen wie Gefallen oder Zufriedenheit. Typische Fragen bzw. Items sind “Ich bin mit dem Training rundum zufrieden”, “Der Veranstaltungsort war geeignet” oder “Der Trainer war kompetent”, “Das Training war nützlich” oder “Das Schwierigkeitsniveau war gerade richtig”.

Auf der Ebene **(2) Lernen** geht es darum, was die Teilnehmer/innen in dem Training gelernt haben. Je nach Inhalt und Zielsetzung des Trainings dürfte es da in der Jugendarbeit vor allem um prozedurales Wissen, bestimmte Fertigkeiten, aber auch Haltungen bzw. Einstellungen handeln. Gerade in der Jugendarbeit wird auch an Haltungen bzw. Einstellungen gearbeitet (Anti-Rassismus-Trainings, Demokratie, Menschenrechte etc.). Bezogen auf MINT ist es ein herausragendes Ziel, Geschlechter-Stereotypen aufzubrechen und das persönliche Wahrnehmungs- und Gestaltungsspektrum zu erweitern.

Auf der Ebene **(3) Verhalten** wird evaluiert, in welchem Ausmaß sich die Lerninhalte eines Trainings bzw. das Erlernte tatsächlich im späteren Verhalten der Teilnehmer/innen niederschlagen, das Erlernte also tatsächlich angewandt wird. Evaluationen auf dieser Ebene 3 zeigen, dass dieser Trainingstransfer nicht automatisch gelingt und das Gelernte im natürlichen Habitat des Teilnehmers nur in Teilen zum Vorschein kommt. Dieses Transferproblem ist seit vielen Jahren Gegenstand der Forschung (u.a. Baldwin & Ford, 1988). Eine Evaluation auf Ebene 3 sollte daher so ausgelegt sein, dass auch die Transferhindernisse erfasst werden. Damit erweitert sich der Fokus, es wird nun der einzelne Teilnehmer als Teil sozialer (und sozio-technische) Systeme gesehen. Dort findet auch die Evaluation statt, und für sie steht im Grundsatz ein breites Spektrum von Methoden und Instrumenten zur Verfügung: Befragung und Beobachtung des jeweiligen Teilnehmers, 360-Grad-Feedback, Logbuch, u.a.m.

Auf der Ebene **(4) Ergebnisse** wird das Training hinsichtlich der Ergebnisse evaluiert, die von der Organisation, die die Teilnahme initiiert hat, als wünschenswert betrachtet werden. Dabei wird es sich vielfach um Ergebnisse handeln, die direkt oder indirekt finanzieller Natur sind. Kirkpatrick (1960b) sieht hier zum Beispiel für einen betrieblichen Kontext die Reduktion von Kosten, Personalfluktuation, Absentismus, Beschwerden; die Steigerung der Qualität und Quantität in der Produktion; und die Verbesserung des Betriebsklimas. Dies lässt sich auf Organisationen und Einrichtungen in der Jugendarbeit übertragen, auch wenn es darum geht die MINT-Kompetenzen und die unternehmerische Haltung von Jugendlichen zu entwickeln. So wird ein Jugendverband wie die Jugendfeuerwehr oder die Jugend im Technischen Hilfswerk ein eminentes Interesse daran haben, dass sich ein Training mit seinen Jugendlichen auch in seiner organisationalen Leistungsfähigkeit niederschlägt.

Kaufman's (und Keller’s) Modell der Lernevaluation

Roger Kaufman und John M. Keller veröffentlichten 1994 ein 5-Ebenen-Evaluationsmodell als Reaktion auf das sehr populäre Modell von Kirkpatrick. Sie greifen kritische Stimmen zur damaligen Praxis der Trainingsevaluation auf, die Evaluation für ein unterbewertetes Unterfangen hielten, das oft unangemessen angewendet wird und manchmal schädliche Folgen hat (Kelle und Möller, 1983). Kaufman und Keller (1994) vermuten, dass es dafür mindestens drei Gründe gibt: (1) Die am häufigsten verwendeten Definitionen und Modelle der Evaluierung sind oft zu eng gefasst, (2) es werden nicht die richtigen Fragen gestellt, die die Evaluierung beantworten soll, und (3) die Beziehung zwischen Zielen und Mitteln wird in den Evaluierungs-, Planungs- und Umsetzungsprozessen nicht deutlich gemacht (S. 371).

Als (eine) Antwort darauf erweitern Kaufman und Keller (1994) den Fokus der Trainingsevaluation und beziehen Ressourcen, die als Befähiger eines Lernprozesses wirken, sowie die gesellschaftlichen Folgen der Trainingsaktivität/-programms mit ein. Ihr Modell umfasst fünf Stufen oder Ebenen: Befähigung (1a)/Reaktion (1b), Erwerb (2), Anwendung (3), organisatorische Ergebnisse (4) und gesellschaftliche Folgen (5).

Auf der Ebene **(1a) Befähigung**, die in den Rezeptionen des Kaufman-Modells überwiegend als Input bezeichnet wird, sind prinzipiell alle Ressourcen evaluierbar, die in den Lehr-/Lernprozess einfließen. Es sind dies vordergründig die zur Verfügung gestellten Lernmaterialien, die technische Ausrüstung, die Qualifikation der Trainer/innen, die Räumlichkeiten. Weiter geschaut sind aber auch die Zugangsvoraussetzungen der einzelnen Lerner und die Zusammensetzung der Gruppe wichtige Inputfaktoren oder Befähiger.

Auf der Ebene **(1b) Reaktion** steht die Lehr-/Lernerfahrung der Teilnehmer/innen im Fokus. Hier werden die Akzeptanz und Effizienz der eingesetzten Methoden, der Mittel und des Ablaufs erfasst und bewertet. Im Anschluss an den ersten Teil wird dieser zweite Teil vielfach auch als Evaluation des Prozesses bezeichnet.

Auf der Ebene **(2) Erwerb** wird die Aneignung der Lerninhalte erfasst und bewertet. Kaufman und Keller sprechen hier von Beherrschung (Mastery) und Kompetenz (Competency), und zwar des einzelnen Lerners wie auch einer kleinen Gruppe. Es geht also darum, ob und in welchem Ausmaß die Teilnehmer/innen die Lehr-/Lernziele erreicht haben. Die Evaluation auf dieser Ebene ist vergleichbar mit der Ebene 2 von Kirkpatrick.

Auf der Ebene **(3) Anwendung** wird erfasst und bewertet, ob und wie gut die Teilnehmer/innen das Gelernte in ihrer Praxis anwenden. Die Evaluation auf dieser Ebene ist vergleichbar mit der Ebene 3 von Kirkpatrick. Im Kaufman-Modell wird diese Ebene (zusammen mit der Ebene 2) auch als Mikro-Ebene bezeichnet.

Auf Ebene **(4) Organisatorische Ergebnisse** wird der Nutzen für die Organisation erfasst und bewertet. Die Evaluation wird aus Sicht der Organisation, die die Teilnahme an der Trainingsmaßnahme initiiert hat, auf der Folie der von ihr gewünschten Ergebnisse vorgenommen. Die Evaluation auf dieser Ebene entspricht der Ebene 4 des Kirkpatrick-Modells. Die Ebene 4 wird im Kaufman-Modell auch als Makro-Ebene bezeichnet.

Die Ebene **(5) Gesellschaftliche Folgen**, die keine Entsprechung im Kirkpatrick-Modell hat, und die von den Autoren auch als Mega-Ebene bezeichnet wird, öffnet den Blick für die Gesellschaft als Ganzes. Wenn denn ein Training einen Beitrag zur Wirksamkeit und zur Effizienz einer Organisation leistet, so stellen die Autoren mit ihrem Evaluationsmodell die Frage in den Raum, ob und in welchem Ausmaß die so gestärkte Organisation (und damit das Training) überhaupt einen wünschenswerten Beitrag zur Wohlfahrt einer Gesellschaft leistet. Auf der Ebene 5 sind Fragen zu Gesundheit, Prosperität, Umwelt, Sicherheit u.a.m. zu stellen und zu beantworten; Fragen, die an Aktualität nichts eingebüßt haben, im Gegenteil.

Das Phillips-ROI-Modell

Phillips (1996,1998) erweitert das Kirkpatrick-Modell um eine fünfte Stufe (oder Ebene), auf der der ROI einer Trainingsmaßnahme evaluiert wird. Der ROI bzw. Return on Investment ist eine Kennzahl aus dem Finanzmanagement, die das Verhältnis zwischen Gewinn und investiertem Kapital angibt, damit ein Indikator für die Rentabilität einer Investition. Phillips überträgt diese Kennzahl auf die Personalentwicklung oder das HRM, mit seinem Evaluationsmodell will er den Wert einer Trainingsinvestition finanziell abbilden (Phillips, 1991). Werden die Modelle von Kaufman (1994) und Phillips (1996,1998) miteinander verglichen, so könnte die jeweilige Erweiterung mit der fünften Ebene (und damit Akzentuierung) kaum unterschiedlicher sein.

Die fünf Ebenen des Phillips-Modells sind (1) Reaktion und geplante Aktion; (2) Lernen; (3) Anwendung am Arbeitsplatz; (4) Geschäftsergebnisse; und (5) Kapitalrendite/ROI.

Auf Ebene **(1) Reaktion** wird wie bei Kirkpatricks Ebene 1 die Zufriedenheit der Teilnehmer/innen mit der Trainingsmaßnahme erfasst und bewertet. Zusätzlich wird im Phillips-ROI-Modell noch gefragt, welche Vorstellungen oder Pläne die Teilnehmer/innen zur Anwendung dessen haben, was sie gelernt haben.

Die Ebene **(2) Lernen** des Phillips-Modells entspricht im Wesentlichen der des Kirkpatrick-Modells. Im Fokus stehen die Lernzuwächse, sie werden auf der Folie der Lehr-/Lernziele erfasst und bewertet.

Auf Ebene **(3) Anwendung** am Arbeitsplatz stehen – wie bei Kirkpatrick (1960a) und Kaufman und Keller (1994) - Verhaltensänderungen am Arbeitsplatz bzw. generalisiert: in der Aufgabenerfüllung. Evaluiert wird, ob und in welchem Ausmaß das Gelernte auch tatsächlich in die Praxis transferiert und dort umgesetzt wird.

Die Evaluation auf Ebene (**4) Geschäftsergebnisse** des Phillips-Modell weist Parallelen zum Kirkpatrick-Modell auf, der Fokus ist allerdings enger gefasst. Auf dieser Ebene 4 geht es darum, ob und inwieweit sich die Teilnahme an dem Training auch tatsächlich in Geschäftsergebnissen einer Organisation niederschlägt. Das kann eine höhere Produktivität sein, weniger Zeitaufwand, weniger Kosten, Qualität, höhere Kundenzufriedenheit und vieles andere mehr. Dass dies auch für Organisationen in der Jugendhilfe interessant sein kann, wurde bereits bei der Besprechung des Kirkpatrick-Modells dargelegt.

Auf der Ebene (5) Kapitalrendite/ROI, der in Phillips Verständnis ultimativen Ebene der Evaluation, wird der monetäre Gewinn der Trainingsteilnahme mit den Kosten verglichen. Für ihn ist die Bestimmung des monetären Kosten-Nutzen-Verhältnisses der notwendige Schritt, um eine Evaluation zu vervollständigen (Phillips, 1998).

CIRO (Context, Input, Reaction, Outcome) Modell

Das CIRO-Modell von Warr, Bird und Rackham (1970) folgt eng einem Prozess guter Planung und Durchführung eines Trainings und setzt entsprechend mit der Evaluation bereits vor der Durchführung der eigentlichen Trainingsmaßnahme ein. Es führt eine neue Ebene ein, die von den Autoren als Kontext-Evaluation bezeichnet wird. Ansonsten zeigen sich viele Parallelitäten zum Kirkpatrick-Modell, zum Teil werden sie allerdings neu und anders strukturiert. Das CIRO-Modell hat vier Ebenen: (1) Kontext, (2) Input (3) Reaktion und (4) Outcome.

Auf der Ebene **(1) Kontext** wird die Qualität der Analyse der aktuellen (und auch zukünftigen) Anwendungssituation, von der ausgehend dann die Trainingsbedürfnisse identifiziert und die Trainingsziele bestimmt werden, evaluiert. Mit diesem erweiterten Fokus wird dem Umstand Rechnung getragen, dass sich schwerwiegende Fehler auf dieser Planungsebene allerspätestens beim ausbleibenden Transfer des Gelernten in die Praxis niederschlagen. Das CIRO-Modell will in diesem Fall die Fehlerursache da identifizieren, wo sie eigentlich liegt.

Auf der Ebene **(2) Input** wird die Qualität der Durchführung der Bildungsmaßnahme inklusive ihrer Konzeption und Planung evaluiert. Das CIRO-Modell richtet seinen Fokus nun darauf, ob und wie Informationen über mögliche Trainingstechniken und -methoden gesammelt, eine Auswahl getroffen, das Trainingsdesign erstellt und das Training durchgeplant wurde. Ein Augenmerk gilt dabei auch der Frage, ob und wie die Ressourcen der Organisation bestmöglich eingesetzt wurden, um die gewünschten Ziele zu erreichen. Dann folgt der eigentliche Input, nämlich das Management und die methodisch-didaktische Umsetzung des Trainings. Alle diese Aspekte liegen im Fokus der Input-Evaluation.

Auf der Ebene **(3) Reaktion** werden die Reaktionen der Teilnehmer*innen erfasst und verarbeitet, um den Prozess zu verbessern. Dies schließt neben den Ansichten über das erlebte Training auch Verbesserungsvorschläge der Teilnehmer/innen mit ein.

Auf der Ebene **(4) Outcome** werden Ergebnisse mit unterschiedlicher Reichweite analysiert: die unmittelbaren Ergebnisse in Form des Lernzuwachses, die Zwischenergebnisse an der Stelle des Transfers in die Praxis und die ultimativen Ergebnisse, die die Organisation dank des Trainings oder des Trainingsprogramms realisieren kann. Damit ähnelt die Evaluation auf dieser Ebene der Evaluation auf den drei Kirkpatrick-Ebenen des Lernens, Verhaltens und der Ergebnisse.

Anwendung bestehender Evaluationsmodelle auf non-formale und informelle Bildung

Non-formale und informelle Bildung bzw. informelles Lernen in einer Organisation der Jugendarbeit ist ein hochkomplexes Geschehen, die Evaluation desselben ist angesichts der Vielzahl der zu kontrollierenden Einflussgrößen so komplex und anspruchsvoll wie der Evaluationsgegenstand selbst. Entsprechend erfordern solche Evaluationen eine Expertise, die in der Regel nur noch bei hochspezialisierten Instituten anzutreffen ist. Sie bedienen sich verschiedener Modelle der Bildungsprogrammevaluation, mit denen zum Teil bereits über

viele Jahrzehnte insbesondere in den Vereinigten Staaten von Amerika in lokalem, regionalen und auch nationalem Zuschnitt Erfahrungen gesammelt wurden.

Die Literatur zur Programmevaluation ist kaum übersehbar, eine angemessene Darstellung der relevanten Modelle oder Ansätze füllt regelmäßig mindestens ein ganzes Buch (u.a. Stufflebeam & Coryn, 2014) oder eine ganze Website (u.a. BetterEvaluation, 2021). Diese seien zur Lektüre empfohlen, wenn es darum geht, eigene Ideen zum Evaluationsansatz in den Gesprächen mit potenziellen externen Auftragnehmern für eine Evaluation zu diskutieren. An dieser Stelle soll nur kurz auf drei Modelle oder Ansätze hingewiesen werden, ohne damit die Bedeutung der vielen anderen Ansätze wie die von Patton (1978) oder Stufflebeam (1971) schmälern zu wollen. Es sind dies die Modelle von Stake (1976), Fetterman (1994) und Cousins und Earl (2005), drei Modelle, die dem Selbstverständnis und dem Auftrag von Organisationen der Jugendarbeit sehr entgegenkommen.

Responsive Evaluation (Stake)

Robert Stake hat seinen Ansatz der responsiven Evaluation als Kontrast zu einem Ansatz entworfen, den er als „übergeordnete Evaluation“ bezeichnete und der sich durch eine formale Bestimmung der Ziele, standardisierte Testung der Studentenleistung, Wertestandards des pädagogischen Personals und Berichte im Stil eines wissenschaftlichen Journals (1976, S. 19) auszeichnete.

Stake stellte dem seinen Ansatz der responsiven Evaluation gegenüber. Sein Ansatz zeichnet sich mehr durch pädagogische Fragen als durch Ziele oder Hypothesen aus, methodisch arbeitet er eher mit direkter und indirekter Beobachtung der Programmteilnahme, er greift die Pluralität der Wertmaßstäbe verschiedener Gruppen auf, nicht nur derjenigen des Lehrpersonals, und zeigt eine kontinuierliche Aufmerksamkeit für die Informationsbedürfnisse der verschiedenen Beteiligten und Anspruchsgruppen (1976, S. 19).

Responsive Evaluation ist nach Stakes‘ Verständnis „keine partizipative Evaluation, aber sie orientiert sich zum Teil an den Anliegen der Interessengruppen, und es ist nicht ungewöhnlich, dass responsives Evaluationsfeedback frühzeitig und während des gesamten Evaluationszeitraums erfolgt.“ (Stake, 2003, S. 66)

Partizipative Evaluation (Cousins und Earl)

Bradley Cousins und Lorna Earl (1992) verstehen ihren Ansatz der partizipativen Evaluation als eine Erweiterung des auf Interessengruppen basierenden Modells, das seinen Schwerpunkt auf eine verbesserte Nutzung der Evaluation durch eine intensivere und breitere Beteiligung der primären Nutzer am angewandten Forschungsprozess legt. Zugleich ist dieser Ansatz ist eingebettet in das Konzept der Professionalisierung (hier von Unterricht) und in das Konzept der Lernenden Organisation bzw. des organisationalen Lernens (u.a. Argyris & Schön, 1978; Argyris, 1993; Huber, 1991; Levitt & March, 1988).

Empowerment-Evaluation (Fetterman)

Der Ansatz der Empowerment-Evaluation (Fetterman, 1994) unterscheidet sich von den bisher angeführten Evaluationsmodellen in mehrfacher Hinsicht. Er ist in gewisser Weise programmatisch und stellt sein Anliegen, Verbesserungen zu ermöglichen und die Selbstbestimmung von Menschen zu fördern, zentral. Dazu ist „ihm jedes Mittel recht“, d.h. auch andere Evaluationskonzepte, ein breites Spektrum von Evaluationstechniken und die kooperativ erarbeiteten Evaluationsergebnisse. Konsequent weist er den Beteiligten und den Evaluatoren auch neue Rollen zu. Die Aufgabe der Evaluatoren ist es, den am Bildungsprogramm Beteiligten die nötigen Verfahren und Instrumente zur Bewertung der Planung, Durchführung und Selbstevaluierung ihres Bildungsprogramms an die Hand zu geben. Idealerweise befähigen sie diese so, dass sie auch über die Evaluation hinaus selbständig in der Lage sind, ihr Bildungsprogramm weiter zu evaluieren und zu verbessern.

Die Realisierung einer Empowerment-Evaluation ist sehr flexibel und mit allen erdenklichen Instrumenten möglich. 10 Prinzipien (Fetterman, 2005), eine sehr einfache Drei-Schritte-Struktur (Fettermann, 2001), die später zu einem Zehn-Schritte-Ansatz (Chinman, Imm und Wandersman, 2004) verfeinert wurde, und ein (nicht abgeschlossener) Fundus an Instrumenten helfen dabei, die Empowerment-Evaluation erfolgreich zu realisieren.

Zusammenfassung

Dieser Beitrag hatte das Ziel, pädagogischen Mitarbeiter/innen in der Jugendarbeit, die vor der Entscheidung stehen, ein Training, ein Bildungsprojekt oder gar ein Bildungsprogramm evaluieren zu wollen oder zu müssen, mit einigen grundlegenden Ausführungen zu den essenziellen Aspekten einer Evaluation zu orientieren. Dazu besprach er zunächst die ethischen Herausforderungen einer Evaluation und die Bedeutung der Zielklärung. Er stellte dann die Vor- und Nachteile der externen und internen Evaluation dar, und danach auf die beiden Evaluationsformen „summativ“ und/oder „formativ“ einzugehen. Die praktische Planung einer Evaluation stand anschließend im Fokus, ergänzt mit einer Darstellung von vier in der non-formalen Bildung anwendbaren Evaluationsmodellen und Hinweisen auf drei Programmevaluationsmodelle, die dem Selbstverständnis und dem Auftrag von Organisationen der Jugendarbeit sehr entgegenkommen.

Literaturverzeichnis

Abele, C. & Erdmann-Kutnevic, S. (2016). *Methodenheft zur Selbstevaluation von internationalen Jugendbegegnungen im Rahmen des Förderprogramms EUROPEANS FOR PEACE der Stiftung »Erinnerung, Verantwortung und Zukunft« (EVZ)* (Überarbeitete und gekürzte Fassung). Verfügbar unter: https://www.stiftung-evz.de/fileadmin/user_upload/EVZ_Uploads/Handlungsfelder/Handeln_fuer_Menschenrechte/Europeans_for_Peace/Handreichung_Selbstevaluation.pdf

Argyris, C. (1993). *Knowledge for action. A guide to overcoming barriers to organizational change* (The Jossey-Bass management series, 1st ed.). San Francisco, Calif.: Jossey-Bass Publishers. Retrieved from http://www.loc.gov/catdir/bios/wiley044/92042861.html

Argyris, C. & Schön, D. A. (1978). *Organizational learning* (Organization development deries). Reading, Mass: Addison-Wesley Pub. Co.

BetterEvaluation. (2013, Mai). *BetterEvaluation. Sharing information to improve evaluation*. Homepage, BetterEvaluation. Verfügbar unter: https://www.betterevaluation.org/en

Bloom, B. S. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives. The classification of educational goals*. Handbook I: Cognitive domain. New York: Longman.

Chinman, M., Pamela Imm & Abraham Wandersman. (2004). *Getting To Outcomes 2004: Promoting Accountability Through Methods and Tools for Planning, Implementation, and Evaluation*. Santa Monica, CA: RAND Corporation. https://doi.org/10.7249/TR101

Council of Europe Youth Work. (2021). *Help: self-assessment. The Council of Europe Youth Work Portfolio,* Council of Europe Youth Work. Verfügbar unter: https://www.coe.int/en/web/youth-portfolio/help-self-assessment

Council of Europe; Conseil de l'Europe. (2018, September). *Self-assessment tool for youth policy*. Brussels. Verfügbar unter: https://rm.coe.int/self-assessment-tool-for-youth-policy-english/16808d76c5

Council of Europe; European Commission. (2021). *Glossary on youth: Training,* Council of Europe; European Commission. Verfügbar unter: https://pjp-eu.coe.int/en/web/youth-partnership/glossary

Cousins, J. B. & Earl, L. M. (Eds.). (2005). *Participatory Evaluation In Education. Studies Of Evaluation Use And Organizational Learning* (Teachers' Library). London: Falmer Press. Retrieved from http://gbv.eblib.com/patron/FullRecord.aspx?p=201009

Cousins, J. B. & Earl, L. M. (2005). The Case for Participatory Evaluation: Theory, Research, Practice. In J. B. Cousins & L. M. Earl (Eds.), *Participatory Evaluation In Education. Studies Of Evaluation Use And Organizational Learning* (Teachers' Library, S. 3–18). London: Falmer Press.

Cronbach, L. J. (1963). Course improvement through evaluation. *Teachers College Record*, *64*(8), 672–683.

Eval-Wiki: Glossar der Evaluation. (2020, 5. November). *Evaluationsplan* (Zeitpunkt der letzten Bearbeitung: 5. November 2020, 19:58 UTC), Eval-Wiki: Glossar der Evaluation. Verfügbar unter: https://eval-wiki.org/glossar/Evaluationsplan

Fetterman, D. M. (1994). Empowerment Evaluation. *Evaluation Practice*, *15*(1), 1–15. https://doi.org/10.1177/109821409401500101

Fetterman, D. M. (2005). A Window into the Heart and Soul of Empowerment Evaluation. Looking through the Lens of Empowerment Evaluation Principles. In D. M. Fetterman & A. Wandersman (Hrsg.), *Empowerment evaluation principles in practice* (1-26). New York: Guilford Press.

Fetterman, D. M. & Wandersman, A. (Hrsg.). (2005). *Empowerment evaluation principles in practice*. New York: Guilford Press.

Guba, E. G. & Lincoln, Y. S. (2001). *Guidelines and Checklist for Constructivist (aka Fourth-Generation).* Western Michigan University. Verfügbar unter: https://wmich.edu/sites/default/files/attachments/u350/2018/const-eval-guba%26lincoln.pdf

Hart, R. A. (1992). *Children's participation. From tokenism to citizenship* (Innocenti essays, no. 4). Florence Italy: UNICEF International Child Development Centre. Verfügbar unter: https://www.unicef-irc.org/publications/pdf/childrens_participation.pdf

Huber, G. P. (1991). Organizational Learning: The Contributing Processes and the Literatures. *Organization Science*, *2*(1), 88–115. Verfügbar unter: http://www.jstor.org/stable/2634941

Kaufman, R. & Keller, J. M. (1994). Levels of evaluation: Beyond Kirkpatrick. *Human Resource Development Quarterly*, *5*(4), 371–380. https://doi.org/10.1002/HRDQ.3920050408

Kellaghan, T. & Stufflebeam, D. L. (Eds.). (2003). *International Handbook of Educational Evaluation* (Kluwer International Handbooks of Education, vol. 9). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-0309-4

Kelle, J. M. & Möller, J. M. (1983). Evaluation of Training: Much Lauded, Seldom Applauded. *IFAC Proceedings Volumes*, *16*(6), 201–209. https://doi.org/10.1016/S1474-6670(17)64365-9

Kirkpatrick, D. L. (1959). Techniques for Evaluation Training Programs. Part 1 - Reaction. *Journal of the American Society of Training Directors*, *13*(11), 21–26.

Kirkpatrick, D. L. (1959). Techniques for Evaluation Training Programs. Part 2 - Learning. *Journal of the American Society of Training Directors*, *13*(12), 21–26.

Kirkpatrick, D. L. (1960). Techniques for Evaluation Training Programs. Part 3 - Behavior. *Journal of the American Society of Training Directors*, *14*(1), 13–18.

Kirkpatrick, D. L. (1960). Techniques for Evaluation Training Programs. Part 4 - Results. *Journal of the American Society of Training Directors*, *14*(2), 28–32.

Kirkpatrick, D. L. & Kirkpatrick, J. D. (2006). *Evaluating training programs. The four levels* (3. ed.). San Francisco, Calif., London: Berrett-Koehler; McGraw-Hill. Retrieved from http://www.gbv.de/dms/bsz/toc/bsz278284698inh.pdf

König, J. (2007). *Einführung in die Selbstevaluation. Ein Leitfaden zur Bewertung der Praxis Sozialer Arbeit* (2., neu überarbeitete Auflage). Freiburg im Breisgau: Lambertus-Verlag. Verfügbar unter: https://www.lambertus.de/assets/adb/94/94d43588099c1578.pdf

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory into Practice*, *41*(4), 212-218. Verfügbar unter: https://www.depauw.edu/files/resources/krathwohl.pdf

Krathwohl, D. R., Bloom, B. S. & Masia, B. B. (1956). *Taxonomy of Educational Objectives. The Classification of Educational Goals*. Handbook II: Affective Domain. New York: David McKay Company.

Levitt, B. & March, J. G. (14). Organizational Learning. *Annual Review of Sociology*, *1988*, 319–340. Verfügbar unter: http://sjbae.pbworks.com/f/levitt_march_1988.pdf

Michaelides, R. & Antonacopoulou, E. (2021, Februar). *Project Management as a Dynamic Collaborative Social Practice. Collaborative Innovation Revisited* (White Paper). PMI Project Management Institute.

Patton, M. Q. (1978). *Utilization-focused evaluation*. Beverly Hills, Calif.: SAGE.

Phillips, J. J. (1996). How Much Is the Training Worth? *Training and Development*, *50*(4), 20–24.

Phillips, J. J. (1998). The Return-on-Investment (ROI) Process: Issues and Trends. *Educational Technology*, *38*(4), 7–14.

Pollitt, C. (1998). Evaluation in Europe. *Evaluation*, *4*(2), 214–224. https://doi.org/10.1177/13563899822208554

SALTO Training and Cooperation Resource Centre. (2021). *YOCOMO self-assessment tool for youth workers,* SALTO Training and Cooperation Resource Centre. Verfügbar unter: https://satool.salto-youth.net/

Scriven, M. [Michael] (1967). The Methodology of Evaluation. In R. W. Tyler, R. M. Gagné & M. Scriven (Hrsg.), *Perspectives of Curriculum Evaluation* (Monograph series on curriculum evaluation, Bd. 1, S. 39–83). Chicago: MacNally. Verfügbar unter: https://files.eric.ed.gov/fulltext/ED030948.pdf

Scriven, M. [Michael]. (1981). *Evaluation Thesaurus* (3rd edition). Verfügbar unter: https://files.eric.ed.gov/fulltext/ED214952.pdf

Simpson, E. J. (1966). *The Classification of Educational Objectives, Psychomotor Domain*. Report Resumes. University of Illinois. Verfügbar unter: https://files.eric.ed.gov/fulltext/ED010368.pdf

Simpson, E. J. (1972). *The Classification of Educational Objectives in the Psychomotor Domain*. Washington, DC: Gryphon House.

Stake, R. E. (1976). A theoretical statement of responsive evaluation. *Studies in Educational Evaluation*, *2*(1), 19–22. https://doi.org/10.1016/0191-491X(76)90004-3

Stake, R. (2003). Responsive Evaluation. In T. Kellaghan & D. L. Stufflebeam (Eds.), *International Handbook of Educational Evaluation* (Kluwer International Handbooks of Education, vol. 9, S. 63–68). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-0309-4_5

Stockmann, R. (2004). *Was ist eine gute Evaluation? Einführung zu Funktionen und Methoden von Evaluationsverfahren* (CEval Arbeitspapiere 9). Saarbrücken:: Centrum für Evaluation, Universität des Saarlandes.

Stufflebeam, D. L. (1971). The Relevance of the CIPP Evaluation Model for Educational Accountability. *Journal of research and development in education*. Verfügbar unter: https://files.eric.ed.gov/fulltext/ED062385.pdf

Stufflebeam, D. L. & Coryn, C. L. S. (2014). *Evaluation Theory, Models, and Applications* (Research methods for the social sciences, Bd. 50, 2nd edition). San Francisco CA: Jossey-Bass.

Tyler, R. W., Gagné, R. M. & Scriven, M. [M.] (Hrsg.). (1967). *Perspectives of Curriculum Evaluation* (Monograph series on curriculum evaluation, Bd. 1). Chicago: MacNally. Verfügbar unter: https://files.eric.ed.gov/fulltext/ED030948.pdf

United Nations Children's Fund (UNICEF): Office of Research - Innocenti. (2013, September). *Ethical Research Involving Children*. Florence, Italy. Verfügbar unter: https://childethics.com/wp-content/uploads/2013/10/ERIC-compendium-approved-digital-web.pdf

Warr, P., Bird, M. & Rackham, N. (1970). *Evaluation of management training. A practical framework, with cases, for evaluating training needs and results*. London: Gower P.

Yarbrough, D. B. & Shula, Lyn M., Hopson, Rodney K., Caruthers, Flora A. (2011). *The program evaluation standards. A guide for evaluators and evaluation users* (3. ed.). Thousand Oaks: SAGE Publ. Retrieved from https://evaluationstandards.org/wp-content/uploads/2019/08/Yarbrough-3e-Rev-Errata-Page.pdf

YouthLink Scotland. (2021). *Self-evaluation and improvement,* YouthLink Scotland. Verfügbar unter: https://www.youthlinkscotland.org/policy-research/outcomes-and-skills/youth-work-outcomes/about-the-youth-work-outcomes/self-evaluation-and-improvement/